

... I am satisfied ...with the awareness and a glimpse of the marvelous structure of the existing world, together with the devoted striving to comprehend a portion, be it ever so tiny, of the Reason that manifests itself in nature.

---Albert Einstein, *The World As I See It*

University of Alberta

Development of monomeric fluorescent proteins and fluorescent protein-based biosensors

by

Hiofan Hoi

A thesis submitted to the Faculty of Graduate Studies and Research
in partial fulfillment of the requirements for the degree of

Doctor of Philosophy

Department of Chemistry

©Hiofan Hoi
Spring 2013
Edmonton, Alberta

Permission is hereby granted to the University of Alberta Libraries to reproduce single copies of this thesis and to lend or sell such copies for private, scholarly or scientific research purposes only. Where the thesis is converted to, or otherwise made available in digital form, the University of Alberta will advise potential users of the thesis of these terms.

The author reserves all other publication and other rights in association with the copyright in the thesis and, except as herein before provided, neither the thesis nor any substantial portion thereof may be printed or otherwise reproduced in any material form whatsoever without the author's prior written permission.

Abstract

Fluorescent protein (FP) technology is now an indispensable tool of biomedical research. Nevertheless, only a few members of the hundreds of existing FPs are generally regarded as the preferred options for most imaging applications. Accordingly, new FPs with ever-improved properties are in demand and worth pursuing, as these efforts may lead to variants that are closer to the “ideal” FP. Also there remain a tremendous number of opportunities for developing FP-based biosensors for probing biological process *in vivo*. This thesis describes our effort on engineering new FPs with improved properties, and further modifying them to create novel biosensors. Directed evolution and semi-rational protein engineering are the main techniques used to develop these new FPs and FP-based biosensors.

The first class of FPs addressed in this thesis are the green-to-red photoconvertible FPs (pcFPs). In an effort to overcome the limitations imposed by the oligomeric structure of natural pcFPs, we created a new monomeric pcFP based on consensus design. Subsequent optimization yielded mClavGR2 and mMaple, two monomeric pcFPs displaying superior performance in folding and maturation, brightness, photoconversion efficiency and photostability. We demonstrate the application of mClavGR2 for dynamic monitoring of protein trafficking. Furthermore, in collaboration with researchers from several other groups,

mMaple was demonstrated to be a multi-model probe that is suitable for use in several conventional and super-resolution fluorescence imaging modalities.

Using mMaple as a template for single-FP biosensor design, we successfully combined the two most important implementations of FPs, the “highlightable” trait and the Ca^{2+} sensing capability, into one construct. Optimization, characterization and live cell imaging of the resulting green-to-red highlightable Ca^{2+} indicators are described.

Another class of FPs that are of interest in this thesis are the true yellow emitting FPs that fill the spectral gap between monomeric greenish-yellow FPs and monomeric orange FPs. By disrupting the inter-subunit interfaces of zFP538, a FP with a distinct three-ring chromophore and an emission maximum at 538 nm, we successfully obtained its monomeric version and named it as mPapaya1. Again, characterization and live cell imaging application of mPapaya1 are described.

Acknowledgements

First and foremost, I would like to thank my supervisor, Professor Robert Campbell, for his ingenious and insightful guidance and suggestions on research, as well as his continuous and considerate help and support throughout the last five years. The qualities, attitudes, knowledge and skills I have learned from him and other people here make my Ph.D. studies at this extremely cold Edmonton worthwhile.

Next I would like to thank all the collaborators throughout the course of my Ph.D. studies: Dr. Nathan Shaner and Dr. Jiwu Wang from Allele Biotechnology; Michael Davidson and Paula Cranfill from the Florida State University and National High Magnetic Field Laboratory; Professor Christopher Cairo from University of Alberta; Dr. Ann McEvoy from University of California Berkeley; and Dr. Tomoki Matsuda and Professor Takeharu Nagai from Osaka University. Without their contributions and help, I would not have accomplished this thesis research.

I would like to express my appreciation to all the professors involved in my supervisory and examing committee: Professor Liang Li, Professor Glen Loppnow, Professor Christopher Cairo and Professor Andrew Simmonds from Department of Cell Biology, as well as Dr. Mark Prescott from Monash University for agreeing to review this thesis.

I would also like to thank all of the Campbell group members, past and present. I am grateful to have you as labmates. Particular thanks go to Huiwang, Andreas, Spencer, Yidan, Yongxin, Yi and Ahmed. My appreciation also goes to many people from the Chemistry Department and outside, Caishun, Blake, Gareth, Xuejun, Lynne, Ben, etc., for their generous favors.

Lastly and as always, I would like to thank my parents and family. People said perseverance is the first important quality to survive a Ph.D. study. If you find any on me, I inherit that from my dear parents.

Table of Contents

Chapter 1 Introduction	1
1.1 Overview and premise.....	2
1.2 Protein engineering	3
1.2.1 General strategies for protein engineering.....	4
1.2.2 Important technologies in protein engineering.....	7
1.3 Fluorescent proteins.....	9
1.3.1 Natural sources.....	9
1.3.2 Three-dimensional structure	10
1.3.3 Chromophore formation.....	12
1.3.4 Photostability	15
1.3.5 Engineered variants of FPs	17
1.3.6 Standard applications of FPs.....	24
1.4 Genetically encoded FP-based biosensors	25
1.4.1 BiFC based biosensors.....	26
1.4.2 FRET based biosensors	27
1.4.3 Single-FP based biosensors.....	30
1.5 The scope of the thesis	33
Chapter 2 Development of monomeric photoconvertible fluorescent proteins	35
2.1 Introduction	36
2.2 Results and discussion.....	39
2.2.1 Guided consensus design of a new pcFP.....	39
2.2.2 Directed evolution with screening for photoconversion.....	40
2.2.3 <i>In vitro</i> characterization of mClavGR2	44
2.2.4 Structure-based comparison of green-to-red pcFPs.....	48
2.2.5 Phenol/phenolate equilibrium and photoconversion	50
2.2.6 Characterization of photostability in live cells	53
2.2.7 Live cell imaging of various mClavGR2 fusion proteins....	54
2.2.8 Imaging protein mobility with mClavGR2	57
2.2.9 Engineering and characterization of mMaple.....	60
2.3 Conclusions.....	63

2.4	Materials and methods	64
2.4.1	General methods and materials.....	64
2.4.2	Plasmid library creation and site-directed mutagenesis....	64
2.4.3	Plasmid library screening.....	65
2.4.4	Protein purification and characterization.....	66
2.4.5	Live cell imaging	68
Chapter 3 Advanced application and characterization of mClavGR2 and mMaple.....		71
3.1	Introduction	72
3.2	Results and discussion.....	78
3.2.1	<i>In vivo</i> folding and maturation of pcFPs.....	78
3.2.2	Reversible photoswitching	80
3.2.3	Photostability study of pcFPs.....	84
3.2.4	mMaple's performance in mammalian cells.....	88
3.3	Conclusions.....	90
3.4	Materials and methods	91
3.4.1	General methods and materials.....	91
3.4.2	Protein purification and characterization.....	91
3.4.3	Live cell imaging	93
Chapter 4 Development of highlightable Ca²⁺ indicators for live cell imaging.....		95
4.1	Introduction	96
4.2	Results and discussion.....	98
4.2.1	Attempted development of a paGECO based on paGFP and green emissive GECOs.....	98
4.2.2	Attempted development of a paGECO based on pamCherry and R-GECO.....	99
4.2.3	Attempted development of a pcGECO based on psOrange and O-GECO	101
4.2.4	Creation of circularly permuted variants of mMaple.....	102
4.2.5	Creation of GR-GECO0.1 and its optimization	104
4.2.6	<i>In vitro</i> characterization of GR-GECOs.....	108
4.2.7	<i>In vivo</i> characterization of GR-GECOs	114

4.3	Conclusions.....	116
4.4	Materials and methods.....	117
4.4.1	General methods and materials.....	117
4.4.2	Generation of circularly permuted variants of mMaple..	117
4.4.3	Plasmid library creation and site-directed mutagenesis..	118
4.4.4	Plasmid library screening.....	119
4.4.5	Protein purification and characterization.....	120
4.4.6	Live cell imaging	122
	Chapter 5 Engineering a monomeric version of yellow fluorescent protein zFP538	124
5.1	Introduction	125
5.2	Results and discussion.....	127
5.2.1	Attempted engineering of monomeric yellow FP with a KYG chromophore through chromophore transplantation	127
5.2.2	Interface disruption of zFP538 and directed evolution	130
5.2.3	Validation of mPapaya1 as a monomer under physiological conditions	138
5.2.4	<i>In vitro</i> characterization of mPapaya1.....	140
5.3	Conclusions.....	141
5.4	Materials and methods.....	142
	Chapter 6 Conclusions and future directions	144
6.1	Summary of the thesis.....	145
6.2	Future works	147
6.3	Closing remark	149
	Bibliography.....	151
	Appendices	173
	Appendix A: Supplementary data for Chapter 2	173
	Appendix B: Supplementary data for Chapter 3	179
	Appendix C: Supplementary data for Chapter 4	181
	Appendix D: Supplementary data for Chapter 5.....	183

List of Tables

Table 2.1 Properties of mClavGR variants and other pcFPs	44
Table 2.2 Photostability of pcFPs in live cells	54
Table 2.3 Properties of mMaple and related variants	61
Table 3.1 Photobleaching rates for pcFP-H2B fusions in live cells.....	88
Table 3.2 Advantages and limitations of mMaple	91
Table 4.1 Properties of mMaple and its circularly permuted variants	104
Table 4.2 Termini library of GR-GECOs	106
Table 4.3 Properties of GR-GECOs	111
Table 4.4 Ca ²⁺ -binding equilibrium of GR-GECOs.....	113
Table 5.1 Properties of mPapaya1 and ZsYellow1	141

List of Figures

Figure 1.1 Schematic representation of FP's application as a fluorescent tag.....	3
Figure 1.2 Schematic representation of the process of directed evolution.	6
Figure 1.3 The structure of avGFP	11
Figure 1.4 Two proposed mechanisms (paths A and B) for chromophore formation in avGFP	13
Figure 1.5 Mechanism for red chromophore formation in DsRed	14
Figure 1.6 Naturally occurring variations of the avGFP chromophore	15
Figure 1.7 The oligomeric structure of wild-type and engineered Anthozoa FPs.....	19
Figure 1.8 Chromophore structures in engineered FPs.....	22
Figure 1.9 Detection of protein-protein interaction by BiFC	27
Figure 1.10 Representative FP-based FRET biosensors	29
Figure 1.11 Circularly permutation of FP	32
Figure 1.12 Cartoon representation of G-CaMP.	33
Figure 2.1 Chromophore structure and UV-induced modifications of green-to-red pcFPs	36
Figure 2.2 Sequence alignment of mClavGR variants and other relevant proteins	41
Figure 2.3 Overview of the directed evolution of mClavGR	42
Figure 2.4 Characterization of the oligomeric structure of mClavGR2, mEos2, and Dendra2 by gel filtration chromatography	43
Figure 2.5 Spectral changes in mClavGR2 upon photoconversion	46
Figure 2.6 Location of positions targeted for semi-saturation mutagenesis.	50

Figure 2.7 Characterization of the pH dependence of the green and red states of mClavGR2	52
Figure 2.8 Photoconversion of unfused mClavGR2 in live cells	55
Figure 2.9 Live cell imaging of mClavGR2 protein fusions	56
Figure 2.10 Photoconversion and imaging of H2B-mClavGR2.....	57
Figure 2.11 Photoconversion of ICAM-1-mClavGR2 on a HeLa cell in contact with a Jurkat cell expressing LFA-1	58
Figure 2.12 Diffusion of ICAM-1-mClavGR2 on the plasma membrane ..	59
Figure 2.13 Gel filtration chromatography of mMaple and mClavGR2	61
Figure 2.14 Spectral characterization and pH titration of mMaple	62
Figure 2.15 Characterization of mClavGR2 and its improved variants	63
Figure 3.1 Lateral resolution of optical microscope	73
Figure 3.2 (f-)PALM/STORM comparison of mMaple, mClavGR2 and mEos2.....	76
Figure 3.3 mMaple has improved <i>in vivo</i> brightness and enables 3D-SIM reconstructions.....	77
Figure 3.4 Expression and maturation of pcFPs in <i>E. coli</i>	79
Figure 3.5 Pseudo native SDS-PAGE gel of the soluble and insoluble fractions of <i>E. coli</i> expressing pcFP-CheWs	80
Figure 3.6 Reversible photoswitching of photoconverted (red) mMaple ..	82
Figure 3.7 (f-)PALM/STORM analysis of purified pcFPs.....	84
Figure 3.8 Widefield imaging photobleaching curves for pcFPs.	86
Figure 3.9 Confocal imaging photobleaching curves for pcFPs.....	87
Figure 3.10 pcFP fusions expressed in mammalian cells	89
Figure 4.1 Photoactivation and Ca ²⁺ response of paR-GECOs.....	101
Figure 4.2 Fluorescent images of <i>E. coli</i> expressing cpmMaple variants on a Petri dish	103

Figure 4.3 Schematic representation of a single FP-based green-to-red photoconvertible Ca ²⁺ indicator	105
Figure 4.4 Sequence alignment of GR-GECOs, G-GECO1.1 and mMaple	107
Figure 4.5 Location of substitutions in GR-GECOs	108
Figure 4.6 Absorbance spectra and pH titration of GR-GECOs.....	109
Figure 4.7 Excitation and emission spectra of GR-GECOs.	110
Figure 4.8 Effect of 460 nm illumination on photoswitching	112
Figure 4.9 Photoconversion profile of GR-GECOs	112
Figure 4.10 Ca ²⁺ titration of GR-GECOs.....	113
Figure 4.11 Stopped-flow kinetic characterization of GR-GECOs	114
Figure 4.12 Representative timecourse traces for HeLa cells expressing cytoplasmic GR-GECO1.1	115
Figure 4.13 Imaging histamine-induced Ca ²⁺ oscillations in highlighted HeLa cells	116
Figure 5.1 Chromophore structures and formation in yellow FPs	126
Figure 5.2 Alignment of mTFP1 and a single unit of tetrameric zFP538	129
Figure 5.3 Cartoon representation of X-ray crystal structure of zFP538	131
Figure 5.4 Zoom views of the inter-unit interfaces in zFP538.....	133
Figure 5.5 Fluorescent image and gel filtration chromatogram of Papaya variants	135
Figure 5.6 Overview of the directed evolution of mPapaya.....	136
Figure 5.7 Sequence alignment of zFP538 and Papaya variants	137
Figure 5.8 Representative widefield fluorescence images of HeLa cells expressing FP-fusions.....	139
Figure 5.9 Spectral profile and pH titration of mPapaya1	140

Abbreviations and Symbols

3D	3-dimensional
avGFP	<i>Aequorea victoria</i> green fluorescent protein
Ala	alanine
Arg	arginine
Asn	asparagine
Asp	aspartate
B-PER	bacterial protein extraction reagent
BiFC	bimolecular fluorescence complementation
BFP	blue fluorescent protein
CaM	calmodulin
HeLa	cervical cancer cell line originating from Henrietta Lacks
cp	circular permutation
CFP	cyan fluorescent protein
Cys	cysteine
dNTPs	deoxynucleotide triphosphates
DNA	deoxyribonucleic acid
DsRed	<i>Discosoma sp.</i> red fluorescent protein
K_d	dissociation constant
DMEM	Dulbecco's modified eagle media
EMCCD	electron-multiplying charge-coupled device
EGFP	enhanced green fluorescent protein
<i>E. coli</i>	<i>Escherichia coli</i>
EGTA	Ethylene glycol tetraacetic acid
FCS	fetal calf serum
FCS	fluorescence correlation spectroscopy
f-PALM	fluorescence photoactivated localization microscopy
FRAP	fluorescence recovery after photobleaching
FP	fluorescent protein
FWHM	full width at half maximum

FRET	Förster resonance energy transfer
GFC	gel filtration chromatography
GECO	genetically encoded calcium optical indicator
Glu	glutamate
Gly	glycine
HHBSS	HEPES-buffered Hank's balanced salt solution
HOMO	highest occupied molecular orbital
His	histidine
IB	inclusion body
ICAM-1	intercellular adhesion molecule 1
Ile	isoleucine
LFA-1	leukocyte function-associated antigen 1
LED	light-emitting diode
LUMO	lowest unoccupied molecular orbital
LB	Luria Bertani
Lys	lysine
mRNA	messenger ribonucleic acid
Met	methionine
mClavGR	monomeric <i>Clavularia</i> -derived green-to-red photoconvertible version 1
MOPS	3-(N-morpholino)propanesulfonic acid
mRFP1	monomeric red fluorescent protein version 1
mTFP1	monomeric teal fluorescent protein version 1
NTA	nitrilotriacetic acid
OD	optical density
Phe	phenylalanine
PMA	phorbol 12-myristate 13-acetate
PBS	phosphate buffered saline
PALM	photoactivated localization microscopy
pcFP	photoconvertible fluorescent protein
PSF	point spread function
PCR	polymerase chain reaction

PDB	protein data bank
POI	protein of interest
PPI	protein-protein interaction
RPR	random priming recombination
ROS	reactive oxygen species
REACH	Resonance Energy-Accepting Chromoprotein
rpm	revolutions per minute
RPMI-1640	Roswell Park Memorial Institute medium-1640
Ser	serine
SDS-PAGE	sodium dodecyl sulfate polyacrylamide gel electrophoresis
StEP	staggered extension process
STED	stimulated emission depletion
STORM	stochastic optical reconstruction microscopy
SIM	structured illumination microscopy
Tyr	tyrosine
Val	valine
WCL	whole cell lysate
YFP	yellow fluorescent protein
ϵ	molar extinction coefficient
Φ	quantum yield

Chapter 1 Introduction¹

¹ A version of the discussion on the chemistry and physics of FPs (i.e. Section 1.3.1 to Section 1.3.5) has been submitted for publication (H. Hoi, Y. Ding, and R. E. Campbell, "FRET with fluorescent proteins". An invited book chapter for *Resonance Energy Transfer*, Ed. I. Medintz and N. Hildebrandt, Wiley VCH Publishing House).

1.1 Overview and premise

Over the past few decades, researchers have made extraordinary progress in understanding biological systems. Of the numerous technologies that propel this process, fluorescent proteins (FPs) stand out as essential tools since they alone provide researchers with the ability to visualize, track and quantify molecular events in living organisms.

FPs are a class of homologous proteins that share the remarkable ability to form a visible wavelength fluorophore from their own amino acid residues. The major advantages of FPs over synthetic dyes as fluorescent tags stem from the fact that FPs enable relatively non-invasive imaging of live organism since they are genetically encodable. That is, following the construction of a recombinant gene consisting of the coding gene of the protein of interest (POI) fused to the coding gene of a FP, introduction of the recombinant gene into cells by transfection allows the fusion protein to be “manufactured” *in situ* using the cells’ protein synthesis machinery. Localization and movement of the POI in the cell becomes observable due to the intrinsic fluorescence of the FP (Figure 1.1). In addition, when fused to specific signal peptides or tissue specific promoters, FPs can be targeted to subcellular compartments or specific tissues of an organism. The discovery and subsequent engineering and optimization of these proteins has provided biologists with a plethora of tools for the imaging of subcellular structures, protein dynamics, metabolite concentrations, and enzyme activities in live cells [1-3]. In recognition of the broad impact of these proteins, the Nobel Prize in Chemistry for 2008 was awarded to Osamu Shimomura, Martin Chalfie, and Roger Y. Tsien, 3 pioneers who made key contributions to the early developments in this field.

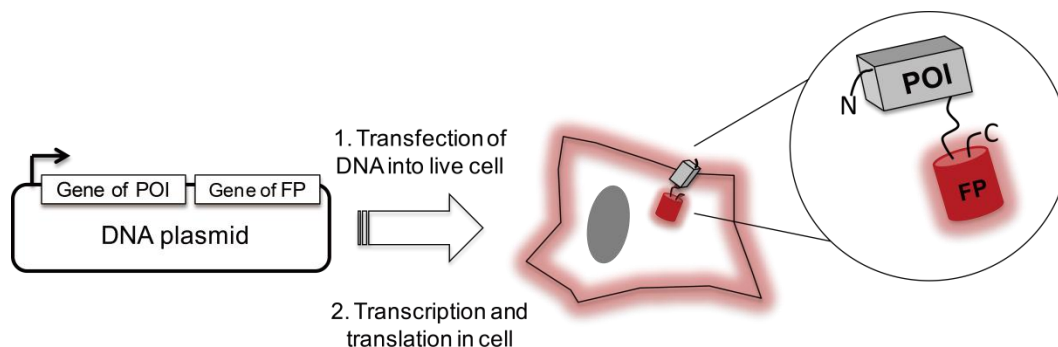


Figure 1.1 Schematic representation of FP’s application as a fluorescent tag. The gene of a red FP is fused to the gene of the protein of interest (POI). After transfection, the fusion protein is expressed in the cell. The fluorescence of the FP marks the location of the POI, in this case, the plasma membrane.

Although FPs were first isolated from natural sources, essentially all FPs currently used in research applications are engineered variants of the wild-type proteins [2, 4, 5]. The word “engineered” here means that the genetic sequence of the FP has been altered in order to improve the properties of the protein. In addition, delicate design and engineering has been used in order to render FPs into satisfactory optical indicators for cellular activities such as metabolite flux and enzymatic activities [3, 6].

In this thesis I will describe our efforts on engineering monomeric FPs and FP-based biosensors. In the remainder of this introductory chapter, I will review several relevant topics including protein engineering, the chemistry and physics of FPs, and FP-based biosensors.

1.2 Protein engineering

Protein engineering is the process of altering properties (stability, photophysical properties, ligand specificity, enzyme activity, etc.) of a target protein by genetically manipulating its coding sequence. Synthesized as linear polypeptides, proteins fold into ordered three-dimensional structures and acquire specific properties and functions as determined by the chemical properties of their primary sequence. The recognition that there exist many examples in which thousands (at least)

of homologous proteins appear to be derived from a common ancestor protein, serves as evidence that proteins are continuously being engineered by nature. While sharing a conserved core structure and often their general function, protein homologues usually have primary sequences substantially different from each other, and may diverge in other properties. The inspiration that we can take from the process of natural protein evolution is that there exists a practically infinite sequence space within which we can manipulate the properties and functions of target proteins by modifying their primary sequence in a laboratory.

1.2.1 General strategies for protein engineering

Two strategies have been widely used in protein engineering: computational design and directed evolution. In computational protein design, researchers seek to create proteins with desired structure and function based on in-depth knowledge of the sequence-structure-function relationship. Proteins, in their native structures, are believed to be in a conformation with global minimum energy. By considering contributions from molecular mechanics (e.g. covalent bond strength, bond angles, torsional potential, Coulomb interactions, van der Waals interactions, solvation effect, etc.), statistical terms derived from solved protein structures, as well as other empirical terms, potential energy functions and computation algorithms are used to find the conformational state with minimum energy for a given sequence, or vice versa (i.e. the sequence that gives the minimum energy for a given structure) [7]. The study of sequence-structure relationship started as early as in 1960s and so far has many fruitful outcomes [8-10]. For instance, one of the landmark works in this area was reported by one of the leading research groups in the area, the Stephen Mayo group at California Institute of Technology. The Mayo group used a computational design algorithm based on physical chemical potential functions and stereochemical constraints to design an entire protein--the buried core, the solvent-exposed surface and the

boundary between core and surface, with a desired structure [11]. Another milestone work came from the David Baker Group at University of Washington. Baker and coworkers successfully designed an artificial protein named as Top7. Top7 not only has novel amino acid sequence, but also has a topology that is not known to occur in nature [12].

The current forefront in computational protein design arguably lies in the area of designing proteins with enzymatic functions. With the enhancement of our ability to perform accurate protein structure modeling, researchers have begun to design proteins with enzymatic or binding functions by optimizing their interaction with other molecules or computed mimics of transition states [13]. Recent breakthrough examples include engineered protein domains that recognize novel target peptide sequences [14], an interface-optimized chimeric endonuclease that can recognize a chimeric substrate site [15], as well as several new enzyme catalysts for reactions that no naturally occurring enzyme has been discovered [16-18].

However, it should be noted that although computational design offers enormous generality to create new proteins with desired activities and traits, formidable challenges remains en route as our scientific knowledge of protein sequence, structure and function is still limited. More accurate energy functions, as well as faster computational algorithms, are required. There is still some way to go before we achieve the goal of being able to routinely and reliably create perfect designer proteins [19].

The natural process of evolution driven by the “survival of the fittest” is the inspiration for the second protein engineering strategy, directed evolution. A typical directed protein evolution strategy involves repeated rounds of: (1) generation of a gene library with random diversification; (2) expression of the gene library in a suitable host such as *E. coli* or yeast; (3) selection or screening for mutants with an improved property; and (4) isolation of “winners” to be used as templates in the next round (Figure

1.2). Generally, these procedures are repeated until satisfactory mutants are acquired. The great power of directed evolution lies in the fact that detailed structural or mechanistic information is not required for the strategy to be successful [20]. On the other hand, directed evolution is highly labor intensive since its success heavily relies on the generation of diversity and effective screening [21]. The Arnold group at California Institute of Technology was among the first to demonstrate the power of directed evolution by evolving a subtilisin E variant which remains active in a highly unnatural environment [22]. Subsequent to that, numerous other successful examples have been reported [23, 24].

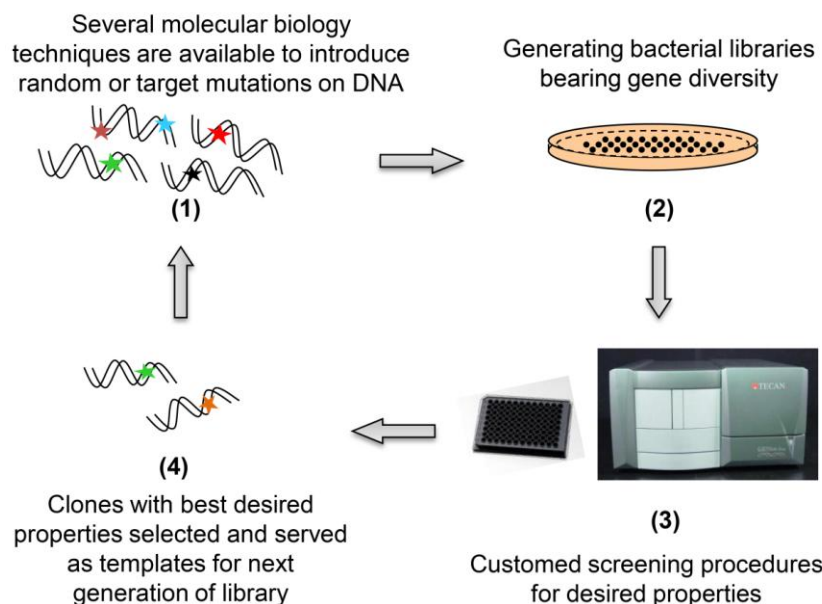


Figure 1.2 Schematic representation of the process of directed evolution.

The two strategies discussed above, directed evolution and computer design, are not mutually exclusive but are complementary. The combined use of both strategies often allows for more efficient protein engineering than either strategy alone [25-27]. For example, it is very likely that only one specific property (e.g. the catalytic activity or the binding affinity) can be computationally optimized. The designed protein, though active, might not be optimized for fundamental properties such as folding or long-term stability. Subsequent optimization of the designer protein by directed

evolution would lead to discovery of variant with both improved catalytic activity and improved folding [18]. Alternatively, a protein sequence activity relationships algorithm could be used to analyze the role of point mutations found in random mutagenesis such that only the beneficial mutations are kept and carried to the new generation [28].

1.2.2 Important technologies in protein engineering

The rapid advance of protein engineering has only been possible due to the revolutionary developments in the area of molecular biology techniques, as well as the universalization of DNA sequencing and synthesis. The most critical molecular cloning techniques in protein engineering lie in the different approaches to create gene diversity. Site-directed mutagenesis, which was first demonstrated by Michael Smith in 1978, is a cornerstone of protein engineering since it allows researchers to change amino acid identities at a precise position of a sequence [29]. The substitution could be a specific amino acid replacement, which is usually chosen using a rational hypothesis. Alternatively it could be randomized to create a library of variants with all 20 of the common amino acids (i.e. saturation mutagenesis) or a subset of predefined amino acid residues (semi-saturation mutagenesis). This technique is particularly useful when knowledge of protein structure and properties is available.

When the structure-function relationship of target protein is unknown, random mutagenesis of the full length of gene is required. Random point mutations are usually introduced by error-prone polymerase chain reaction (PCR) in which the error rate of the DNA replication process is intentionally increased. This is realized by using a polymerase with low fidelity (e.g. *Taq* polymerase). Furthermore, the mutation frequency can be controlled by addition of Mn^{2+} [30], unbalancing the available dNTPs pool in the reaction, and varying the number of cycles in the PCR process [31].

However, since most point mutations are deleterious or neutral, a suitably low random point mutation rate is generally preferred [32]. As a result, accumulation of beneficial mutations and evolution of a desired property is relatively slow. Fortunately, the development of *in vitro* gene recombination is one approach that has helped to overcome this drawback of random mutagenesis. In *in vitro* gene recombination methods (often referred to as DNA shuffling), genes of proteins are randomly fragmented and the fragments reassembled to form hybrids through a self-priming polymerase reaction [33, 34]. Gene templates for the shuffling reaction are usually improved mutants isolated from random libraries. By doing gene shuffling, the benefits from different templates can be collectively combined into one variant. In addition to “hit-shuffling”, homologous protein genes from the same structural family but different species can also serve as templates for DNA shuffling. For instance, a cephalosporinase variant with up to 540-fold increased activity was successfully evolved from four wild-type cephalosporinases (58-82% identity at the DNA sequence level) [35]. Stemmer’s initial protocol of DNA shuffling required a relatively tedious step of fragmentation and reassembling. Technically improved gene recombination methods such as random priming recombination (RPR) [36], and staggered extension process (StEP) were developed to simplify the procedure and accumulate point mutations more efficiently [37].

The screening procedure is another crucial step for successful directed evolution, since it is widely accepted that “you get what you screen for” [38]. Current molecular biology techniques can readily create mutant libraries with the size of 10^4 - 10^7 variants. However, without effective screening procedures, such a massive sequence space ends up as a waste and beneficial mutations are overlooked. Fortunately, by designing library screening strategies where the desired protein properties can be correlated to readable signals such as luminescence, binding affinity or a

growth advantage, very large libraries become amenable to high throughput screening [38, 39].

1.3 Fluorescent proteins

1.3.1 Natural sources

The first FP to be identified in nature was the *Aequorea victoria* jellyfish green FP (avGFP). In 1962, Osamu Shimomura reported the discovery of this protein while undertaking the purification of the bioluminescent protein Aequorin from the same jellyfish [40]. In addition to being the first FP to be discovered, avGFP was also the first to be applied in live cell imaging experiments in the mid-1990s [41]. AvGFP has since been joined by a vast number of variants and homologues from other species that differ in various spectroscopic properties, yet share a common 3-dimensional (3D) protein fold and the same core chromophore structure [42].

Until 1999, only a few bioluminescent marine organisms including *Aequorea* jellyfish and the sea pansy *Renilla reniformis* [43], were widely known to be sources of naturally occurring FPs. In both of these examples, the GFP is serving as an energy transfer acceptor for a bioluminescent enzyme. This energy transfer process effectively shifted the blue luminescence to longer wavelength green fluorescence. Since the first FPs had been discovered in bioluminescent organisms, it was generally thought that additional FP homologues would most likely be discovered in other bioluminescent animals. It therefore came as a bit of a surprise to the community at large when Matz *et al.* reported that nonbioluminescent reef corals of the Anthozoa class are a source of avGFP homologues [42]. Interestingly and rewardingly, these Anthozoa FPs come in variety of fluorescent hues that span the cyan to red regions of the visible spectrum [44, 45]. Of these Anthozoa FPs, the one that has been the subject of the greatest interest and further engineering is the *Discosoma sp.* red FP (DsRed) [42]. In recent years, avGFP homologues

have also been found outside the Cnidaria phylum, in crustacean copepods and chordate amphioxus [46-48]. Together, avGFP, DsRed, and just a handful of the 100s of other Anthozoan FPs, have served as the progenitors of the numerous engineered variants that have been optimized for particular live cell imaging applications and are widely available within the research community (see Section 1.3.5).

The biological roles of FPs in reef corals is still very much an open question, but it is likely that they have some role in providing photoprotection to photosynthetic symbiotic algae in shallow waters [49], and in aiding identification by co-habiting species of fish [50]. It is interesting to note that FPs are the only known examples where the color of an organism can be encoded by a single gene [44].

1.3.2 Three-dimensional structure

To facilitate further explanation of FP protein and chromophore structure, it is helpful to focus on the archetypical avGFP. As the first FP to be cloned [51] and have its X-ray crystal structure determined [52, 53], avGFP has been subjected to much more extensive scrutiny than any of its homologues. AvGFP is a protein of 238 amino acids with an approximate molecular mass of 27 kDa [51]. Following transcription of the 714 base pair gene that encodes avGFP, and subsequent translation of the mRNA to the polypeptide, the protein spontaneously folds into a cylindrical 11-stranded antiparallel β -sheet structure that is often referred to as either a β -barrel or β -can (Figure 1.3) [52, 53]. Running along the central long axis of this cylinder is a segment of the protein, composed of residues 57-70, that is in a primarily helical conformation. Located within this helical region and close to the geometric centre of the cylindrical structure are the three residues, Ser65-Try66-Gly67 that undergo the autogenic post-translation modifications to form the conjugated system of the fluorophore (Figure 1.4) [54, 55]. For the remainder of this chapter, the conjugated system of a FP will be generally referred to as a chromophore.

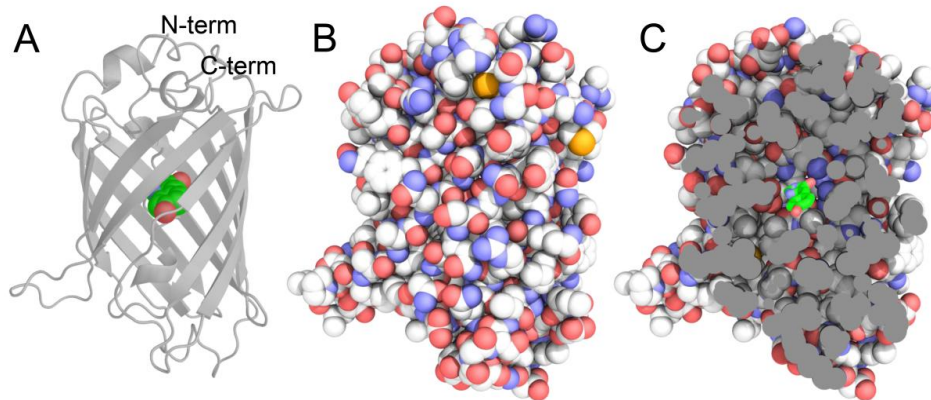


Figure 1.3 The structure of avGFP. (A) A cartoon representation of avGFP (Protein Data Bank (PDB) ID 1EMA) with the chromophore shown in space-filling representation with carbon atoms colored green, nitrogen atoms colored blue, and oxygen atoms colored red [52]. (B) An all-atom representation of avGFP with carbon atoms colored white, sulfur atoms colored orange, and oxygen and nitrogen colored as in (A). (C) The same representation as (B) but with the protein sliced through the middle to reveal the chromophore which is shown in a stick representation and colored as in (A). This slice does bisect the chromophore, but the whole chromophore is shown for the sake of clarity.

The autogenic formation of the chromophore is absolutely dependent on the complete and proper folding of the β -barrel. Fortunately, the folding of the nascent polypeptide into the complete β -barrel structure is generally a robust process (particularly in variants that have been engineered for efficient folding at 37 °C) and relatively dramatic modifications of the polypeptide chain can be tolerated [56, 57]. For example, it is possible to make circular permutation (cp) variants of avGFP in which the original N- and C- termini are linked by several inserted residues, and new N- and C- termini created at a position elsewhere in the protein [58-60]. Once a FP is correctly folded and the chromophore has formed, it is quite stable and, as far as proteins go, quite resistant to a variety of environmental insults. However, when exposed to elevated temperatures, relatively high concentrations of chemical denaturants, strongly acidic conditions, or very high pressure, unfolding of the protein will occur [61]. Since the integrity of the β -barrel is critical for keeping the chromophore isolated from bulk

solvent and thus maintaining the high fluorescence quantum yield, even partial unfolding of a FP can lead to quenching of its fluorescence.

1.3.3 Chromophore formation

The defining feature of the class of FPs is their ability to promote the autogenic post-translational modification of their own amino acid sequence to form the conjugated system of a visible wavelength fluorophore. Given the intense interest in these proteins and the vast number of researchers who have used them in their research, the precise details of the chromophore formation reactions have been surprisingly elusive. A long-standing proposal for the mechanism of the fluorophore forming reactions invoked a three-step reaction in which the three-residue sequence of 65-66-67 underwent a cyclization of the main chain followed by loss of a molecule of water and finally an oxidation of the C_{α} - C_{β} bond of Tyr66 with molecular oxygen (path A in Figure 1.4) [62]. This series of steps is known as the cyclization-dehydration-oxidation mechanism. In recent years, accumulating evidence has eroded support for the cyclization-dehydration-oxidation mechanism and it now appears more likely that the mechanism precedes via cyclization-oxidation-dehydration as shown in path B in Figure 1.4 [63]. A recent study used single molecule imaging to determine that the time constant for chromophore formation in an avGFP variant was about 5 minutes [64].

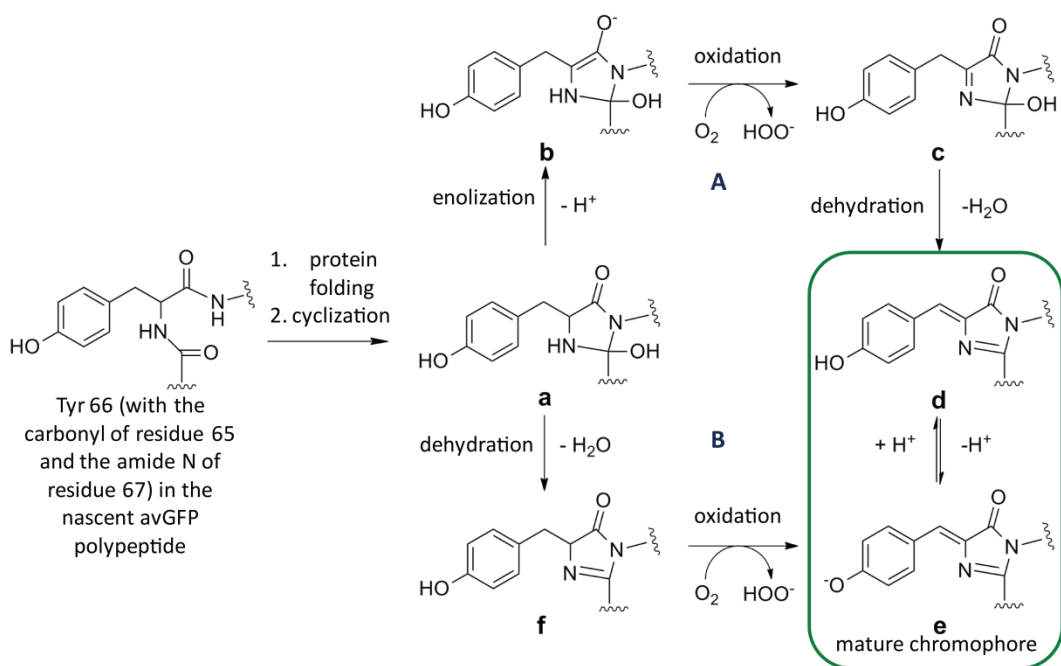


Figure 1.4 Two proposed mechanisms (paths A and B) for chromophore formation in avGFP. Path A represents the original cyclization-dehydration-oxidation mechanism [62]. Recent years have seen growing support for the cyclization-oxidation-dehydration mechanism of path B. The exact nature of the species that undergoes the oxidation remains unclear, and the carbonyl tautomer is another reasonable possibility (i.e. no enolization step). The mechanism shown here was adapted from a review by Wachter [65]. Wavy lines represent points of attachment to the remainder of the polypeptide chain.

Relative to avGFP, there has been far fewer reports of effort to decipher the mechanism of chromophore formation in DsRed. However, the latest research in this area has suggested the chromophore formation in both green and red FPs can be explained using a single mechanistic scheme that involves a branched pathway [66]. Indeed, path B in Figure 1.4 represents one route along this branched pathway; the route that leads to a green fluorescent chromophore. The branch point in the pathway is the first oxidation product (intermediate **c** in Figure 1.4, equivalent to intermediate **a** in Figure 1.5). In contrast to avGFP, the first oxidation product in DsRed is subject to loss of OH⁻ and a second irreversible oxidation that commits the protein to form a red, rather than green, fluorophore (Figure 1.5). Strack *et al.* obtained support for this mechanism

through a very elegant series of experiments in which the ratio of green to red chromophore was altered depending on isotope effects and the concentration of oxygen [66]. The branched pathway mechanism also provides an explanation for the origin of a blue fluorescent intermediate species that has been trapped and structurally characterized by X-ray crystallography [67].

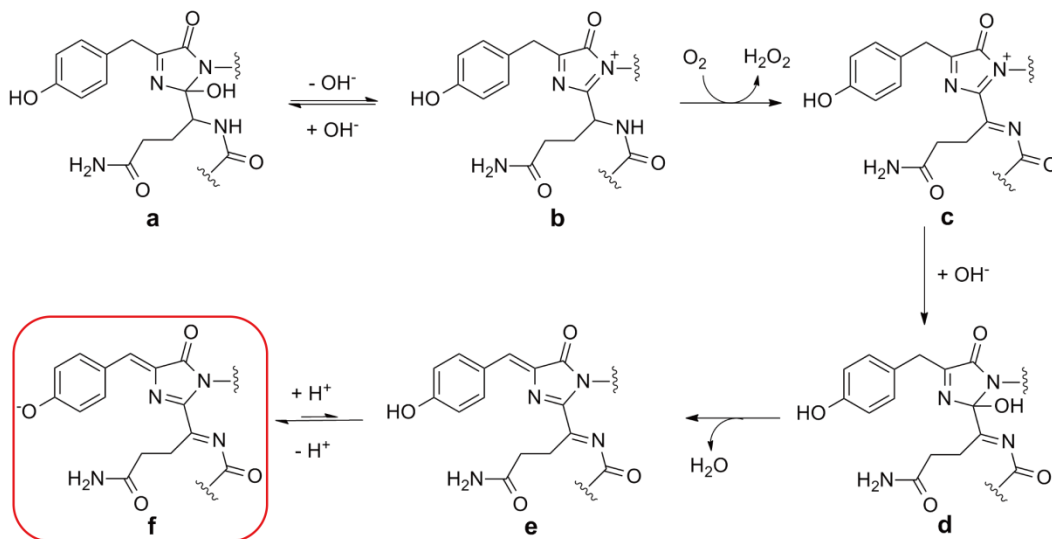


Figure 1.5 Mechanism for red chromophore formation in DsRed. A branch point occurs during the avGFP chromophore formation pathway at intermediate **c** of Figure 1.4 (equivalent to intermediate **a**). An extra oxidation step leads to formation of the blue fluorescent intermediate **c** and subsequent oxidation and dehydration give the red fluorescing chromophore.

All FPs are members of a single protein superfamily and thus share a common folded structure, namely the β -can structure as shown in Figure 1.3. As expected for structurally similar proteins, the chromophores also share a common structure. The diversity of colors that has been observed in naturally occurring FPs is attributed to several relatively modest elaborations of the avGFP chromophore (Figure 1.6). As a general rule, substitution of the first amino acid residue in the chromophore-forming tripeptide, together with assistance from residues in close proximity, allows additional modification on the avGFP chromophore or the DsRed chromophore. For instance, the orange emission in Kusabira Orange, an

FP with a Cys65-Tyr66-Gly67 chromophore forming triad, is proposed to be the result of nucleophilic attack by the thiol side chain of Cys65 on the acylimine moiety of a red intermediate (Figure 1.6C) [68, 69].

It is also possible for the chromophore to exist in either the *Z* (often referred to as *cis*) or the *E* (often referred to as *trans*) stereoisomer as determined by the configuration of the double bond between the C_α-C_β bond of Tyr66. The chromophores represented in Figure 1.6 are shown as the more common *Z* stereoisomer. Several naturally occurring FPs have been shown to have chromophores that exist primarily as the *E* stereoisomer [70, 71], and yet other FPs have been shown to undergo photoinduced isomerization between the *Z* and *E* isomers [72]. Formation of the green-to-red photoconvertible Kaede-type chromophore (Figure 1.6D) and the yellow chromophore in *Zoanthus* yellow FP (Figure 1.6B) will be discussed in detail in Chapter 2 and Chapter 5, respectively.

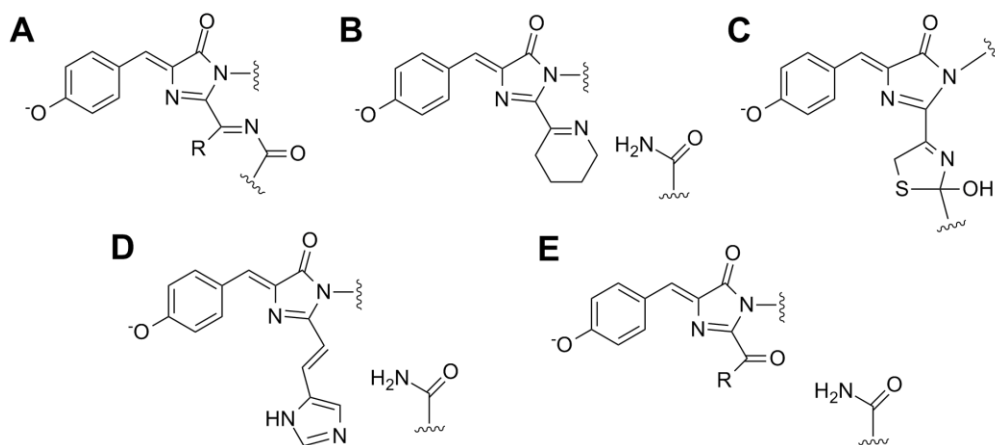


Figure 1.6 Naturally occurring variations of the avGFP chromophore. (A) The red fluorescent DsRed chromophore [73]. (B) The *Zoanthus* yellow FP chromophore [74]. (C) The *Fungia concinna* Kusabira Orange FP chromophore [68, 69]. (D) *Trachyphyllia geoffroyi* Kaede red FP chromophore [75, 76]. (E) The *Anemonia sulcata* kindling FP chromophore [77, 78].

1.3.4 Photostability

Fluorescence imaging, regardless of whether it uses arc lamp or laser excitation, involves exposing cells to very high light intensities. This high

light intensity is necessary for achieving a strong fluorescence signal from the FPs present in a cell, but it also simultaneously causes the destruction of the same molecules it is being used to detect. FPs, as is typical for organic chromophores, undergo excited state reactions that can lead to loss of fluorescence (photobleaching) or alteration of their fluorescent spectrum or spectral properties (photoconversion). Practically speaking, these processes are generally considered undesirable since they contribute to a loss of fluorescence signal with respect to time. However, there are now a subset of FP imaging techniques, such as fluorescence recovery after photobleaching (FRAP), that do take advantage of the ability of FPs to be photobleached [79].

The tendency to photobleach upon illumination is probably the most notorious property of FPs though, generally speaking, the rates are not so dissimilar from those of synthetic dye-based fluorophores. For example, relative to fluorescein, EGFP, mCherry, mCitrine, and mOrange are 23.9, 13.1, 6.7, and 1.2-fold more photostable, respectively [4]. The mechanism of photobleaching is still poorly understood but it is generally thought that a primary mechanism is reaction with reactive oxygen species (ROS) [80]. Indeed, it has recently been shown the reaction with hydroxyl radical leads to a decrease in CFP intensity and lifetime [81]. Interestingly, some FPs display reversible photobleaching in which fluorescence diminishes because the chromophore is driven into a dark state by the excitation light. One of the best-characterized examples of the reversible formation of a dark state is the photoinduced isomerization of a FP chromophore from the *Z* to the *E* state [82]. Since the process is reversible, the FP chromophore can undergo the *E* to *Z* isomerization and reform the normal (and fluorescent) isomer of the chromophore if left in the dark. It is likely that this process is responsible for the reversible on-off blinking that has been observed to occur in avGFP [83] over a timescale of several seconds [82].

Most of the recently developed FPs or commonly-used FPs have had their photostability thoroughly characterized at one or two different illumination intensities that correspond to typical widefield and confocal imaging conditions [84, 85]. The photostability under the relevant experimental conditions should be a key consideration for researchers when choosing FPs for their studies. However, it is now generally recognized that the relationship between the intensity of the illumination light and the FPs' photostability is not necessarily linear. This is an ongoing concern, since practically no FPs have been thoroughly characterized in terms of photobleaching at more than one or two different light intensities.

1.3.5 Engineered variants of FPs

The modifications of a FP's properties can range from the subtle to the extreme. One of the earliest and least obtrusive modifications of FPs is the complete synthesis of the gene with altered codon usage and no change in the actual amino acid sequence of the protein [86]. Moving along the spectrum towards more dramatic alterations, we find some modified variants that only differ from their wild-type parents by a few point mutations that improve the spectral properties or improve the protein folding at 37 °C [56, 57]. Moving further along the spectrum towards higher degrees of alteration, there are some commonly used variants that harbor dozens of mutations relative to their wild-type progenitor [87]. These mutations can have dramatic effects on the properties of the FP, including changes in the oligomeric structure or extreme shifts in the excitation and emission wavelengths. At the extreme end of the modification scale, we find some variants that represent "consensus" designs of FPs and can no longer be said to be direct descendants of a single natural protein ancestor [88, 89]. These proteins differ enough from any protein found in nature that they might as well represent a protein from a hypothetical (non-existent) species that has undergone divergent evolution over millions of

years. A recent development in the area of FP engineering is to use computational methods to design libraries of FP variants optimized for a particular property [27].

1.3.5.1 Engineered monomeric FPs

With the discovery of FP color variants in Anthozoa [42], many new opportunities for engineering of improved or hue-shifted variants became available to protein engineers. However, the Anthozoa FPs do suffer from a critical drawback that had to be addressed by protein engineering before further elaboration of the palette of variants would be justified. The drawback is that the Anthozoa FPs are obligate tetrameric proteins in which four copies of the β -barrel must come together to form the native form of the protein (Figure 1.7). Such tetrameric structure ensured that any protein fused to Anthozoa FPs would become tetrameric itself and thus its normal localization and intracellular dynamics were certain to be perturbed. For this reason, Anthozoa FPs are generally considered incompatible for fluorescent labeling and production of their monomeric variants is prerequisite for further engineering and wide application.

The first example of an engineered monomeric version of an Anthozoa FP was the monomeric version of DsRed known as mRFP1 [87]. In tetrameric DsRed, each protomer, denoted as A, B, C, and D as in Figure 1.7, is in contact with two neighboring ones such that there exist two types of subunit interfaces [90, 91]. In order to make mRFP1, surface residues in DsRed involved in protein-protein contact were genetically modified such that the interface is destabilized [87]. For a tetrameric protein, this is a two-step process since there are two different protein-protein interfaces. Disrupting one interface produces a dimeric FP and disrupting both of them produces a monomeric FP (Figure 1.7). Due to their obligate tetrameric structure, converting Anthozoan FPs to monomeric variants has proven to be a non-trivial undertaking [87, 92]. A total of 33 amino acid substitutions (out of 225 amino acid in DsRed) were introduced during the

engineering of mRFP1: 13 of the substitutions were responsible for the interface breaking-down; while the other 20 were introduced to rescue the lost brightness [87].

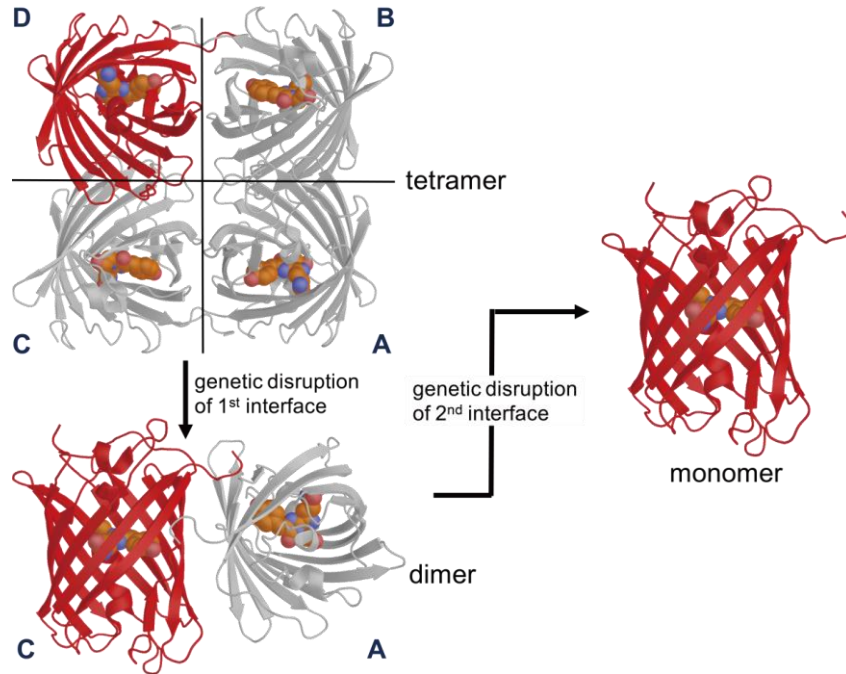


Figure 1.7 The oligomeric structure of wild-type and engineered Anthozoa FPs. Wild-type Anthozoa-derived FPs are homotetramers. However, dimeric and monomeric versions of these proteins can be engineered by genetically changing residues in the protein-protein interfaces to amino acids that are incompatible with the formation of the tetramer or dimer structure, respectively.

Proteins that have been engineered to be monomeric typically have a lower case “m” preceding their names (e.g. mRFP1), whereas dimers have a “d” (e.g. dTomato), and tandem dimer have “td” (e.g. tdTomato) [87]. Tandem dimers are made by fusing together two copies of the FP gene, one after another, with a short linker sequence. In tandem dimer proteins, the formerly intermolecular dimer interactions are now intramolecular interactions and the protein is effectively monomeric, but twice the size of a true monomer. With two copies of chromophore, tandem dimers display enhanced brightness and photostability.

Sequence and X-ray crystal structure alignment revealed that tetrameric structure in the Anthozoa FPs are highly conserved. That is, they all have two different protein-protein interfaces with each one duplicated in the tetramer due to symmetry [71, 74, 93, 94]. This general similarity of the interfaces means that a strategy similar to that used to produce mRFP1 can be applied to other Anthozoan FPs, leading to an expansion of the toolbox of available colors of monomeric FPs for use in imaging applications [92, 95]. Interestingly, some Anthozoa FP seems to be naturally more amenable to being “monomerized” due to intersubunit interfaces that are much less extended than those found in DsRed. As a result, a few amino acid mutations (3 in some cases) are enough to generate functional monomeric variants [68, 96-99]. However, achieving a protein that remains monomeric at high protein concentrations might require much more effort [100-102]. The monomerized variants have, themselves, served as the templates for further modification and expansion of the FP palette (see Section 1.3.5.2) [103, 104].

It is worthwhile to mention that avGFP, which is generally monomeric at low concentrations, has also been reported to form a dimer with a K_d of 110 μM [105]. This K_d is high enough that negligible amounts of dimer species will be present under normal imaging conditions where intracellular concentrations are kept at or below the low μM range. Fortunately, if this weak dimeric tendency is problematic for a particular application (e.g. imaging of membrane-targeted proteins), it can be readily eliminated by the single substitution Ala206Lys [105].

1.3.5.2 Engineered hue-shifted FPs

As the spectral properties of FPs are mainly determined by the combined influences of the chemical structure of their chromophore and the chromophore environment, modification of either leads to FPs with new spectral properties. There is now a wide selection of engineered FPs available for multicolor imaging and other applications, with emission

maxima ranging from the blue to the far-red region of the visible spectrum. In an extraordinary example of multi-color imaging, six subcellular structures were labeled with six FPs with different color-hues respectively. The six emission signals were efficiently separated by spectral unmixing when excited with a single laser line and provided a clear look at the interactions of different structures in a living cell [95].

Roger Y. Tsien and coworkers pioneered the engineering of hue-shifted FPs and produced some of the first color variants of avGFP and mRFP1 [62, 103]. Most notable are the initial versions of the cyan (CFP), blue (BFP), and yellow (YFP) avGFP variants [1]. Blue-shifting the emission of avGFP was majorly achieved by direct modification on the chromophore by replacing Tyr66 with other aromatic amino acid such as tryptophan (e.g. CFP, Figure 1.8A), histidine (e.g. BFP, Figure 1.8C), and phenylalanine (e.g. a blue FP known as Sirius Figure 1.8B) [106], all of which effectively reduce the conjugated system of the chromophore. In contrast, red-shifting avGFP is realized by modification of the chromophore environment. Specifically, Thr203 was substituted with tyrosine to introduce a π - π stacking between its phenol ring and that of Tyr66 (Figure 1.8H). A blue fluorescent avGFP derivative, known as mKalama1, with a tyrosine derived chromophore has also been engineered. In mKalama1, excitation of the higher energy protonated state of the chromophore (species **d** in Figure 1.4) produces blue fluorescence rather than green fluorescence [107].

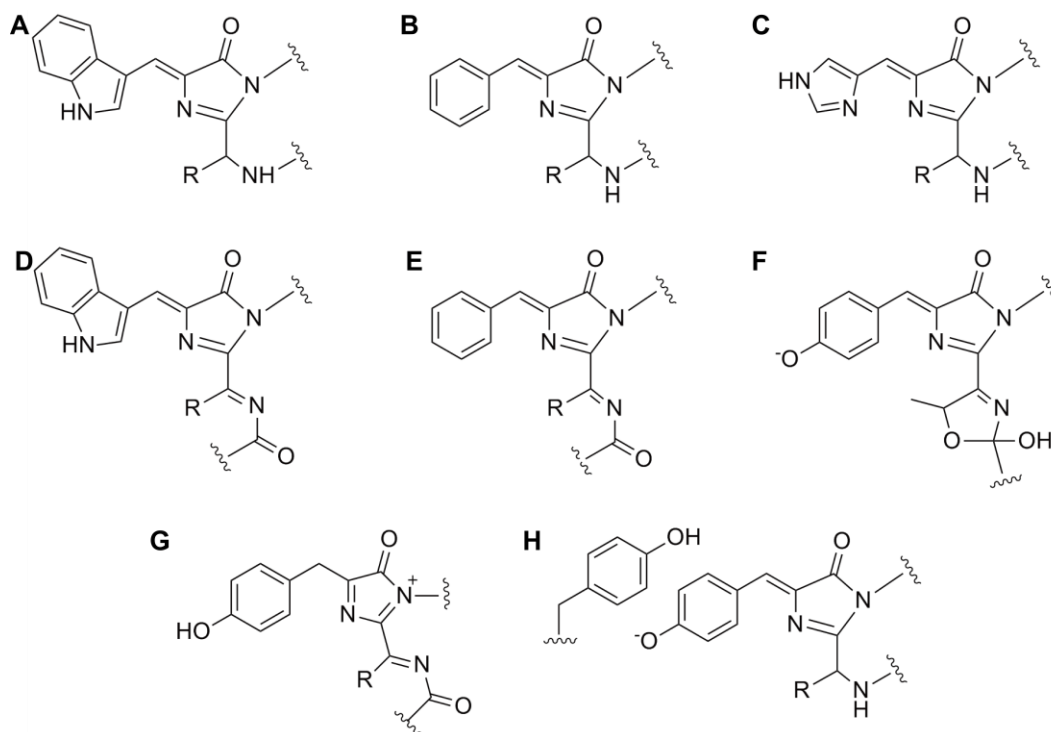


Figure 1.8 Chromophore structures in engineered FPs. (A) Variants of avGFP with the Tyr66Trp mutation [62] are commonly known as cyan FPs (CFPs) [1]. (B) Sirius is an optimized variant of the Tyr66Phe mutant [1] of avGFP and emits blue fluorescence [106]. (C) The Tyr66His variant [62] of avGFP is commonly known as blue FP (BFP) [108]. (D) The tryptophan analogue of the DsRed chromophore is found in the yellow fluorescent mHoneydew variant [103]. (E) The phenylalanine analogue of the DsRed chromophore is found in the blue fluorescent mBlueberry variant [107]. (F) The chromophore of the orange fluorescent variant known as mOrange is derived from a DsRed-type chromophore but contains a third ring due to addition of the adjacent threonine hydroxyl to the main chain acylimine moiety [109]. (G) A proposed structure for the blue fluorescent chromophore of mTagBFP [66, 67, 110]. This structure is also proposed to be an intermediate in the DsRed chromophore formation mechanism (Figure 1.5) [66].

Hue-shifted derivatives of mRFP1 are mostly known as the mFruit series, with the color of fruit giving a rough indication of the color of the FP [103]. These include blue-shifted variants that have covalently altered chromophore structures such as mHoneydew (Figure 1.8D), mBlueberry (Figure 1.8E) [107] and mOrange (Figure 1.8F), and red-shifted variants such as mCherry and mPlum [111]. The bathochromic effect is the result

of chromophore environment change, usually an accumulated contribution from hydrogen bonding and electronic interactions between the chromophore and chemical groups in the residues in close proximity [109, 112].

Other FP progenitors that have been subjected to intense engineering include eqFP578 and its homologue eqFP611 from *Entacmaea quadricolor* sea anemone [113, 114]. These are promising templates for engineering far-red or infrared FPs since they are among the most red-shifted natural FPs discovered to date. A redder color hue is preferred in live cell imaging and deep tissue imaging due to minimal autofluorescence and better light penetration at the spectral range known as the near infrared optical window (i.e. 600 nm to 1200 nm). In addition, low energy excitation light causes less phototoxicity. The covalent structures of the chromophores in eqFP578 and eqFP611 derivatives such as TagRFP [114], mRuby [115], mKate2 [116], mNeptune [102], eqFP650 [117], and eqFP670 [117] remain identical to that of DsRed. The red-shifting in their emission spectra is again a combination of synergistic or additive effect of a large number of amino acid substitutions, some of which are remote from the chromophore. Many of these substitutions were discovered during directed evolution efforts and their roles remain poorly understood [102, 118]. A blue-shifted variant of TagRFP was also created by trapping the red chromophore maturation at its blue intermediate state (i.e. intermediate **c** in Figure 1.5) [110]. Accordingly, the resulting FP known as mTagBFP has a chromophore that is composed entirely of double bonds in the main chain of the protein (Figure 1.8) [67].

1.3.5.3 Other engineered FPs

While much of the earliest interests in FP engineering focused on removing detrimental properties such as oligomeric tendency, thermosensitivity, inefficient maturation and pH sensitivity, or introducing new color-hues, many recent efforts have focused on improving

photophysical properties such as enhanced brightness, better photostability, and better quantum yield [84, 119, 120]. In addition, some atypical properties customized for specific applications are becoming the goal of some FP engineering efforts. For instance, several long Stokes shift FPs with different emission maxima were developed for multiparameter imaging [85, 95, 121, 122]. There is a growing trend to use long Stokes shift FPs in multiparameter imaging applications, since they can provide a spectrally distinct color options without changing the excitation filter [122, 123]. Another atypical class of FPs is the growing number of photo-modulatable FPs that were engineered for protein dynamics study and super-resolution imaging (also see Chapter 2 for detailed discussion). A YFP variant with high extinction coefficient but low quantum yield known as the Resonance Energy-Accepting Chromoprotein (REACH) was developed to be a Förster resonance energy transfer (FRET) acceptor for green fluorescent FP donors [124].

1.3.6 Standard applications of FPs

As briefly discussed in Section 1.1, FPs are widely used as genetically encoded fluorescent tags to study the target protein's fate in a cell such as their localization, movement, and turnover (Figure 1.1). FPs have also traditionally served as reporters for transcriptional regulation. When the gene of FP is placed under control of a specific gene promoter, switching the promoter "on" leads to expression of the FP gene, with fluorescent intensity serving as an indication of the activity of the promoter or other transcription factors involved in the transcription activation.

While conventional microscopy set-ups are used to study the distribution of FP-labeled proteins at their steady state over time, more complicated biochemical behavior of proteins can be studied using special imaging techniques such as fluorescence recovery after photobleaching (FRAP) and fluorescence correlation spectroscopy (FCS), both of which have been broadly applied to study protein kinetics [79, 125]. In a typical FRAP

experiment, a small volume of FP-labeled proteins are photobleached by high intensity illumination. Fluorescence in the photobleached area will recover due to diffusion of unbleached proteins into the bleached area. Analysis of the recovery kinetics can provide key biochemical insights such as whether the protein is freely diffusible or bound to a scaffold, the fraction of immobilized protein, and the association/dissociation dynamics, etc. [79]. The recent emergence of photo-modulatable (photoactivable, photoswitchable and photoconvertible) FPs provides an alternative approach for FRAP [126, 127]. Instead of being photobleached into “invisible” non-fluorescent states, photomodulatable FPs can change their spectral properties when illuminated with light of a certain wavelength. For example, FPs can change from non-fluorescent to fluorescent or from green-fluorescent to red-fluorescent, upon irradiation with light of specific wavelength [128]. Therefore, protein dynamics can be directly monitored in the new “highlighted” spectral channel. In a FCS setup, fluorescence intensity of a very small volume of solution is monitored. By recording and correlating the fluctuations in fluorescence intensity, researchers can gain quantitative information on the diffusion behavior, local concentration, complex formation and other photochemical or photophysical properties of FP-labeled protein in a live cell [125].

A more intricate application of FPs is in the context of genetically encoded biosensors, which will be discussed in more detail in next section.

1.4 Genetically encoded FP-based biosensors

Biosensors generally refer to analytical devices that are composed of two elements: a biological or biologically-derived recognition element which can interact with specific target; and a transducer element which can convert the recognition event into a measurable signal. The term genetically encoded biosensors specifically refer to biosensors whose transducer element is also a biomolecule, thereby the whole sensor can

be synthesized and assembled in live cell [6]. FPs are excellent candidates as transducers since formation of the fluorophore is autocatalytic and fluorescence can be readily detected with high sensitivity. Several approaches have been developed to convert FPs into effective transducers for a variety of recognition events. Based on the design strategy, genetically encoded FP-based biosensors can be categorized into three main classes: bimolecular fluorescence complementation (BiFC) based biosensors, Förster energy transfer (FRET) based biosensors, and single FP based biosensors.

1.4.1 BiFC based biosensors

Complementation is the reconstitution of two protein fragments into a functional whole protein. In BiFC based biosensors for detection of protein-protein interactions, the FP is genetically split into two parts and each part is fused to one of the potentially interacting protein pair. The FP fragments are essentially non-fluorescent and only minimum complementation happens at low concentration. When there is protein-protein interaction (PPI) between the two protein partners, the FP fragments are brought into close proximity and reassemble into a functional whole FP (Figure 1.9). Since Regan and coworkers successfully demonstrated the reconstitution of split avGFP fragments in 2000 [129], FPs with alternative color-hues such as cyan, yellow, red and far-red, have proved to be capable of undergoing fragment complementation [130, 131]. Provided certain knowledge of the structure and conformation of the proteins of interest, construction of such BiFC based biosensors is relatively straightforward and optimization is usually not required. Nonetheless, the complementation is irreversible, which is useful for detection of transient PPI but becomes problematic when disassembly of the protein complex is also of interest [132]. This irreversibility can be overcome by FRET-based biosensor or the recently developed fluorogenic biosensors based on FP heterodimerization [133].

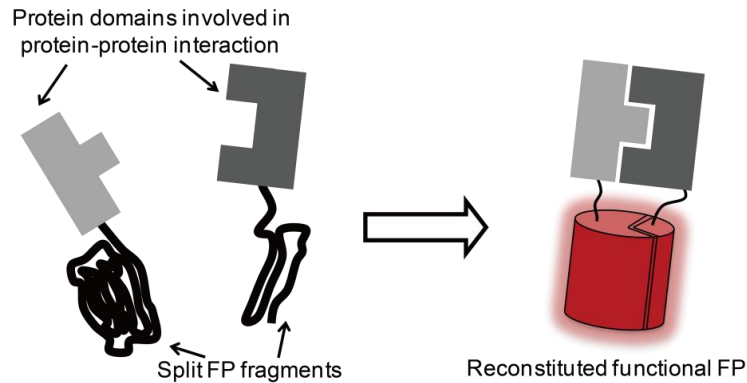


Figure 1.9 Detection of protein-protein interaction by BiFC. Two potentially interacting proteins are fused to the two split FP fragments respectively. Interaction (including external triggered interaction) between the two protein domains would bring the FP fragments into close proximity, which then reconstitute into a whole FP to give fluorescence as signal. The complementation is effectively irreversible.

1.4.2 FRET based biosensors

FRET is the radiationless energy transfer through dipole-dipole interaction between two chromophores that have compatible energy levels and are close in distance (<10 nm) [134, 135]. In FRET based biosensors, the design rationale is to couple a specific PPI or modification of a protein or a peptide to the efficiency of the energy transfer between the donor FP and the acceptor FP [6]. Using appropriate design approaches, a variety of FP FRET-based biosensors for ions or small molecules, ligand-induced PPI or protein conformation change, protease activities, as well as enzyme activities, have been developed [136]. Some typical designs are illustrated in Figure 1.10. The intermolecular design for PPI detection is similar to that of BiFC, with the donor FP and acceptor FP fused to potential interacting protein partners respectively. FRET occurs when the two protein partners form a complex (Figure 1.10A). However, unlike BiFC, FRET between the FPs does not prevent the dissociation of their protein partners.

Another widely used design for ligand induced PPI or protein conformation change is to incorporate both FPs and both protein partners

in one single construct to allow for intramolecular FRET (Figure 1.10B). The most famous FRET based biosensors are probably the Ca^{2+} indicators known as cameleons. In a cameleon, two different colors of FP are genetically fused by a polypeptide region composed of the Ca^{2+} binding protein calmodulin and its binding peptide M13 [137-139]. In the absence of Ca^{2+} , the calmodulin/M13 region is largely unstructured and the two FPs are spatially further away from each other. Therefore FRET is relatively inefficient and excitation of the donor results in more fluorescence from the donor and less from the acceptor. Upon Ca^{2+} binding, calmodulin wraps around M13 to form a globular complex, bringing the two appended FPs closer together. As a result, FRET increases and thus excitation of the donor gives rise to less emission from the donor and more emission from the acceptor. Swapping calmodulin and M13 with protein domains responsive to other ions or small molecules has led to generation of a variety of FRET-based biosensors for analytes such as Zn^{2+} [140], cAMP [141], and various sugars [142].

PPI or conformational change resulted from post-translational modification by enzymes is also exploited to design FRET based biosensors for enzyme activity (Figure 1.10D). Typical examples for this design include an array of kinase biosensors such as AKAR1-3 for protein kinase A [143-145], BKAR for protein kinase B [146], and Src reporter for Src [147], among others. Similar to the cameleons, in these kinase biosensors a substrate peptide and its binding domain are sandwiched by two paired FPs. Phosphorylation of the substrate peptide triggers intramolecular complex formation and changes the distance or relative orientation between the two FPs, thereby changing the FRET efficiency [145].

The last class of biosensors based on FRET principle that will be discussed here are biosensors for proteolytic enzymes (Figure 1.10C). Construction of such biosensors is straightforward; a peptide containing the protease substrate-sequence is flanked with a FRET pair of FPs.

Proteolysis at the substrate sequence lead to separation of the two FPs and thus a loss of FRET signal. Examples in this class include sensors for caspase-3 [85], poliovirus 2A protease [148] and hepatitis C virus NS3/4A protease [149]. Such biosensors are attracting an increasing amount of attention, as they are useful in developing high-throughput screening assays for protease inhibitors of clinical importance [148, 150].

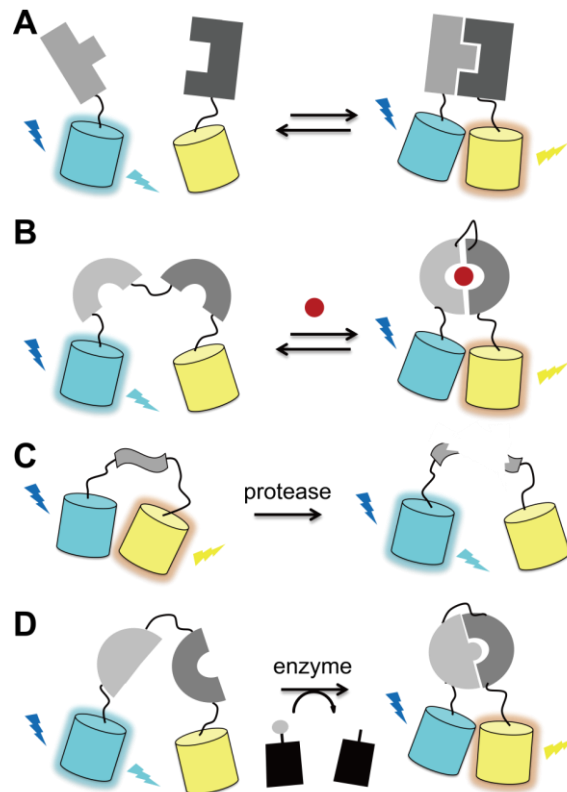


Figure 1.10 Representative FP-based FRET biosensors. (A) Intermolecular biosensors for PPI. Unlike BiFC based biosensors, the FRET pair does not interrupt disassembly of the protein complex formed by their fusion partners. (B) Intramolecular biosensors for ligand (represented by red dot) induced PPI. In some cases, the interacting protein domain might undergo conformational change in order to form the complex. (C) Protease biosensors where the two FPs are linked by a protease-labile sequence. (D) Enzyme activity biosensors. Protein complex formed (or conformation changed) upon installation of specific chemical functionality to the substrate domain by the enzyme. In all illustrations, the cyan can represents the FP donor, while the yellow can represents the FP acceptor. For design A, B and D, although FRET is result from PPI between two protein domains, it could also be the result of conformational change of a single protein domain induced by the presence of ligand or enzyme activity.

Although an impressive palette of FPs is now available, only a small number of them have been found particularly useful for FRET based biosensors [3]. In addition to the generally desirable physical and chemical properties such as robust folding, high brightness and high photostability, an ideal FRET FP pair should have a large spectral overlap between the emission of the donor and the absorbance of the acceptor, while having a small spectral bleed-through between the emission of the donor and the emission of the acceptor [6]. The cyan-yellow pair is the prevailing choice for a number of good reasons: they share excellent spectral overlap, and cyan and yellow FPs with high and comparable brightness are available. Other pairs such as blue-green [108], cyan-orange [68], green-red [151], orange-red [152] and red-far red [153], have also been reported.

The FRET approach is a highly robust and versatile strategy for biosensor development. In addition, accurate quantitative measurement is readily achievable since the measurement is inherently ratiometric and thus only minimally affected by variables such as biosensor concentration and autofluorescence. Nonetheless, a serious disadvantage of FRET-based biosensors is that it is usually difficult to render the biosensors highly sensitive and the optimization is a rather empirical process [136, 154].

1.4.3 Single-FP based biosensors

As suggested in the name, the third design for genetically encoded FP-based biosensors engages only one FP. The spectral signal, should it be fluorescence intensity change or spectra shift of the FP, is rendered responsive to changes in the external environments such as the concentration of analytes. A straightforward designing strategy for single-FP based biosensors is to take advantage of the inherent biosensing capability of the FPs. For instance, pH sensitivity and halide ion sensitivity are found in some YFP variants and regarded as detrimental for many fluorescent tagging and FRET applications [155]. However, this intrinsic

sensitivity also means they can serve as indicators for pH or halide ions [156-158]. Structural and spectral analysis revealed that the key Tyr203 mutation in YFPs is also responsible for introducing a halide ion-binding site in the vicinity of the chromophore [159]. Anion binding at this cavity promotes protonation of the chromophore and leads to suppressed fluorescence. Based on the sensing mechanism, improved halide ion probes with high affinity for cell transport studies and with better selectivity of I⁻ over Cl⁻ have been engineered [160, 161]. In addition to engineering the inside of the FP barrel, subtle modification at the surface of FP can also lead to biosensors. For example, biosensors for disulfide bond formation and redox potential were created by installing cysteines on the surface of avGFP variants [162, 163]. In both cases, conformational changes associated with disulfide bond formation lead to distortion of the beta-barrel and reorganization of the local environment of the chromophore [162, 163].

More dramatic engineering of the FP by genetic incorporation of an extrinsic recognition element into the FP barrel can lead to the generation of tailored biosensors. The extrinsic recognition domain must be fused to the termini or inserted at the loop area of the FP barrel such that it would not disrupt the folding of the overall structure [58]. However, in order for FP to work as an effective transducer, the recognition domain must also be physically close to the chromophore. Unfortunately for this application, the FP chromophore is well protected in the center of the barrel (Figure 1.3C) and the termini are located relatively distant from the chromophore at the top of the barrel (Figure 1.3A). The solution is to generate circularly permuted variants of FPs with new N- and C- terminal close to the chromophore [164, 165]. As briefly mentioned in Section 1.3.2, circularly permuted FP (cpFP) is generated by linking the original N- and C-termini with a short polypeptide spacer and introducing new N- and C-terminal at a position elsewhere in the protein (Figure 1.11) [58-60]. Ligand induced conformational changes at the extrinsic recognition

element, which is fused to the N- or C- terminal of the cpFP, will result in alteration of the chromophore environment and thereby its fluorescence.

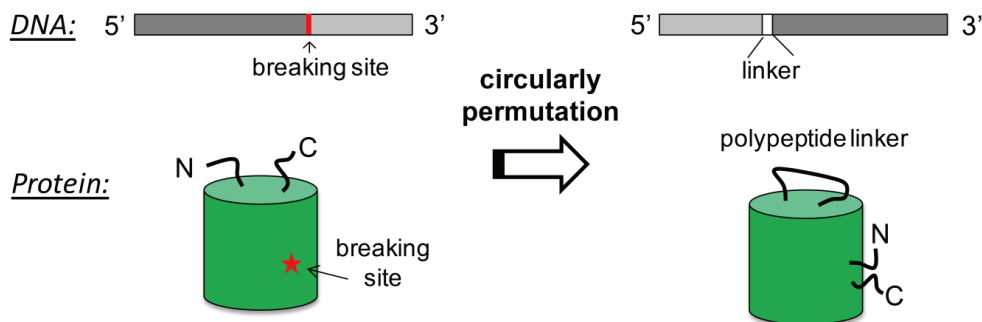


Figure 1.11 Circularly permutation of FP. The original N- and C- termini of the FP are connected with a linker and new N- and C- termini are created at a position close to the chromophore. The 3D structure of cpFP remain the same though the primary sequence is permuted.

The prototypical biosensor of this design is the Ca^{2+} sensor known as G-CaMP (Figure 1.12) [165]. In G-CaMP, M13 and calmodulin (CaM) were fused to the N- and C- terminal of cpEGFP149-144 (numbers denoting the GFP numbering of the new N- and C- termini). Structural studies reveals that in its Ca^{2+} free state, the fluorescence is quenched because the chromophore is freely accessible to bulk solvent [166, 167]. In the presence of Ca^{2+} , CaM wraps around M13 to form a compact structure that blocks the solvent access and restores the chromophore to its fluorescing state. In addition to solvent blocking, a serendipitous hydrogen-bond interaction between Arg377 in the calmodulin and the chromophore in the Ca^{2+} bound state contributes to the improved dynamic range in G-CaMP2. Based on the same rationale, a red Ca^{2+} biosensor known as R-GECO1 was created by replacing the cpEGFP in G-GECO1.1 (an improved G-CaMP variant) with a circularly permuted variant of mApple (i.e. cpmApple146-145) followed by directed evolution [168].

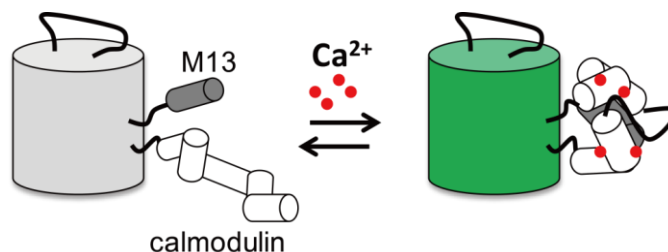


Figure 1.12 Cartoon representation of G-CaMP. In the presence of Ca^{2+} , CaM wraps around M13 to form a compact complex that blocks the solvent access to the chromophore and restores the fluorescence.

1.5 The scope of the thesis

It is now widely acknowledged that FP technology is an established and trusted tool [3]. However, for better or worse, each of the existing FPs has its own set of advantages and limitations. Accordingly, new FPs with improved properties are still worth pursuing, as these efforts may lead to variants that are closer to the “ideal” FP. Also there remain a tremendous number of opportunities for developing FP-based biosensors for probing biological process *in vivo*. This thesis describes our effort on engineering new FPs with improved properties, and further modifying them to create novel biosensors. Directed evolution and semi-rational protein engineering are the main techniques used to develop these new FPs and FP-based biosensors.

Chapter 2 and Chapter 3 describe the development of new monomeric green-to-red photoconvertible FPs (pcFPs). Using a well-characterized monomeric FP known as mTFP1 as template, we designed a new monomeric pcFP that is designated as mClavGR1. Subsequent engineering lead to mClavGR2 and mMaple, both with improved properties such as folding, brightness, photostability and photoconversion efficiency. Engineering and characterization of the new pcFPs are described in Chapter 2. Their utility for single molecule localization-based superresolution imaging, which was performed by collaborators, is presented in the introduction section of Chapter 3. The remaining of the Chapter 3 investigates the reasons for the new pcFPs’ exceptional

performance in live cell imaging. Development of mClavGR2 has been described in a full paper published in *Journal of Molecular Biology* [89]. Our paper on the development and application of mMaple has been published in *PLoS One* [169].

In Chapter 4, mMaple is further developed into a photoconvertible Ca^{2+} biosensor based on the single-FP design. We also tried to convert the cpFP module of existing single-FP based Ca^{2+} biosensors into photoactivable or photoconvertible variants. Nonetheless, only the former strategy ultimately proved to be successful. Characterization and live cell imaging applications of the new biosensors is also described. A manuscript describing this work has been accepted by *Journal of the American Chemical Society* for publication (Hoi, H.; Matsuda, T.; Nagai, T.; Campbell, R. E. "Highlightable Ca^{2+} indicators for live cell imaging", *Journal of the American Chemical Society* 2012, DOI: 10.1021/ja310184a.).

Chapter 5 describes our efforts to engineer a monomeric yellow FP using zFP538 as a template. zFP538 has a unique three-ring chromophore structure formed by the tripeptide Lys-Tyr-Gly and emits maximally at 538 nm. The true yellow emission will fill the current spectral gap between monomeric yellow FP ($\lambda_{\text{em}} \sim 527$ nm) and monomeric orange FP ($\lambda_{\text{em}} \sim 565$ nm) and may potentially serve as better FRET donor for red FPs. By breaking the intersubunit interfaces one by one, we successfully acquired a monomeric variant of zFP538 and named it as mPapaya1. Again, characterization and live cell imaging application is described. This work is in preparation for submission.

In the last Chapter, a summary of this thesis work and an outlook for the future developments in the field of FP engineering is provided.

Chapter 2 Development of monomeric photoconvertible fluorescent proteins²

²The research involving the engineering and characterization of mClavGR2 described in this chapter has been published in *Journal of Molecular Biology* [89], while the research involving the engineering of mMaple has been published in *PLoS One* [169].

2.1 Introduction

In recent years a handful of *Aequorea* green fluorescent protein homologues cloned from Anthozoan organisms have been reported to undergo irreversible photoconversion from a green fluorescent to a red fluorescent species upon illumination with light of approximately 400 nm. To date, the naturally occurring photoconvertible FPs (pcFPs) that have received the most attention are Kaede from coral *Trachyphyllia geoffroyi* [75], EosFP from stony coral *Lobophyllia hemprichii* [98], and Dendra from octocoral *Dendronephthya sp.* [97]. It has also been demonstrated that a naturally occurring non-pcFP can be engineered to be a pcFP. Specifically, the pcFP known as KikGR was engineered from the green fluorescent variant, KikG of the coral *Favia favaus*. [170].

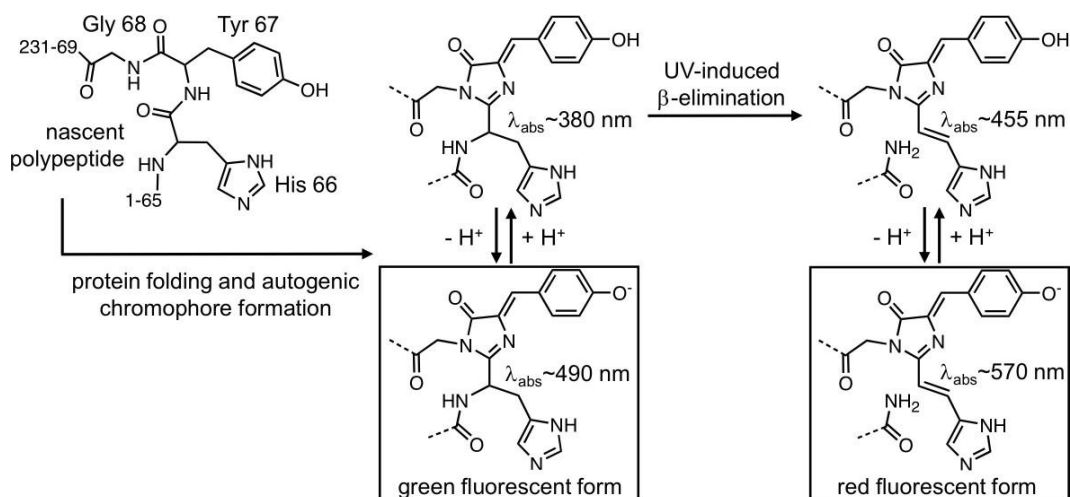


Figure 2.1 Chromophore structure and UV-induced modifications of green-to-red pcFPs [76]. Residue numbers for the nascent polypeptide are for the mTFP1-derived variants described in this work. Dashed lines represent connections to the polypeptide chain of the protein.

All green-to-red pcFPs characterized to date share a common His-Tyr-Gly-derived chromophore structure and a common photoconversion mechanism (Figure 2.1) [76]. The newly synthesized protein first folds into the characteristic β-barrel structure of the avGFP superfamily [52], and then undergoes the steps of post-translational modification that lead to the

formation of a green fluorescent chromophore with a conjugated system identical to that of avGFP chromophore. The chromophore can exist in either its neutral phenol form or the anionic phenolate form (Figure 2.1). Exactly where the equilibrium between these two forms lies is dependent on the local microenvironment of the chromophore (as determined by the amino acid substitutions in close proximity to it) and the pH of the solution. The green-to-red pcFPs are distinguished from their non-photoconvertible brethren by the respective consequences of exciting the neutral form which absorbs most strongly at ~400 nm. In wild-type avGFP the excited state of the neutral species undergoes excited state proton transfer to form the anionic species, which then emits green fluorescence [171]. In the case of the green-to-red pcFPs, excitation of the neutral form leads to a break of the polypeptide chain through the effective β -elimination of the residue that immediately precedes the chromophore-forming His-Tyr-Gly tripeptide [76, 172]. This elimination reaction results in the installation of a new double bond between C_{α} and C_{β} of the His residue, placing the side chain imidazole into conjugation with the remainder of the avGFP-type chromophore. This extended conjugation decreases the HOMO-LUMO gap and shifts the emission into the orange-to-red region of the visible spectrum. Photoconversion (i.e. a light-induced change in excitation or emission wavelength maxima) via alternate mechanisms has been observed in other color classes of FP [173, 174].

FP photoconversion allows researchers to “highlight” a subpopulation of proteins within a cell or tissue through spatially defined illumination with a specific wavelength of light [2]. The subsequent dynamics of the highlighted protein can be followed due to its distinct fluorescence excitation and emission [79, 126]. Optical highlighting can also be achieved using FPs that undergo so-called photoactivation (irreversible conversion from a non-fluorescent to a fluorescent species upon illumination) [175] and photoswitching (reversible conversion between a non-fluorescent and a fluorescent species upon illumination) [176]. A more

recent application of these types of proteins, that has generated excitement in the cell biology community, is in so-called super-resolution fluorescence imaging [177]. This application also involves highlighting a sub-population of protein molecules, but instead of being in a spatially defined location, they are sparsely distributed throughout the whole sample to be imaged. The sample is then imaged to reveal the precise locations of the point sources of fluorescence, each of which corresponds to a single molecule of a FP. The sample is bleached and the process is repeated. Many repetitions of this imaging protocol can produce images in which structures with dimension of 20~30 nm can be resolved [178].

Kaede, the first example of a FP that can undergo an irreversible green-to-red photoconversion upon illumination with UV light, was initially described by Miyawaki and coworkers in 2002 [75]. Unfortunately, the range of potential applications for Kaede remains limited by the fact that it is an obligate tetramer [179], and no monomeric variants have been reported. Unlike monomeric FPs, tetrameric FPs are generally detrimental to the proper trafficking and localization of recombinant fusion proteins [87]. To produce a monomeric green-to-red pcFP, the same workers appear to have had more success with engineered variants of the tetrameric KikGR FP which is substantially brighter and more efficiently photoconverted than Kaede [170]. A monomeric version of KikGR, known as mKikGR, was reported [180].

The two other green-to-red pcFPs, EosFP and Dendra, have been subjected to protein engineering to convert the wild-type tetramers into monomers [97, 98]. However, it is apparent that the monomeric version of EosFP retains a tendency to form dimers at higher concentrations [100]. The monomeric variant of EosFP, known as mEos, was created through the introduction of two point mutations that disrupted the protein-protein interfaces of the tetrameric species [98]. Expression of mEos at temperatures greater than 30 °C is problematic, but an effectively

monomeric tandem dimer variant does express well at 37 °C [181]. The poor expression of mEos at 37 °C has been overcome with the engineering of mEos2 through targeted substitution of residues with solvent-exposed side chains [100]. Although mEos2 has been reported to retain some propensity for dimer formation, this property does not appear to have adverse effects on the subcellular targeting of a variety of fusion proteins [100].

The growing number of reports of optimized pcFPs reflects the growing demands on these proteins with respect to their enabling role in some popular cell biology applications. In particular, there is a continued need for bright and monomeric photoconvertible proteins since these tools can enable the highest precision super-resolution fluorescence imaging [100, 180]. In an effort to create a new monomeric pcFP with favorable properties, we embarked on a strategy which is distinct from that previously employed. Rather than start with a tetrameric photoconvertible protein and engineer it to be monomeric, we sought to start with a well-characterized monomeric coral-derived FP and engineer it to be a green-to-red pcFP. Precedent that it would be possible to convert a non-photoconvertible FP into a pcFP is provided by the fact that tetrameric KikG was converted to KikGR through the use of protein design [170]. Our starting template is a monomeric version of *Clavularia* sp. cyan FP (cFP484) known as mTFP1 (monomeric teal FP 1) [92].

2.2 Results and discussion

2.2.1 Guided consensus design of a new pcFP

Design of the first generation of the new pcFP designated as mClavGR1 is credited to Nathan C. Shaner (Allele Biotechnology).

The new pcFP based on the mTFP1 template was designed using a strategy analogous to that previously used to design a consensus FP

based on a monomeric Azami green template [88]. The known pcFPs, including EosFP [98], Dendra2 [97], mKikGR [180], and Kaede [75], were aligned to find the consensus at each amino acid position (Figure 2.2). Amino acids of >50% consensus were maintained in the designed protein. At positions with no clear consensus the corresponding residue in mTFP1 was used. Residues of mTFP1 that had been substituted during the conversion of the wild type cFP484 tetramer into a monomer, were maintained in the design of the consensus protein. The designed gene was synthesized and, when expressed in *E. coli*, was found to exhibit green fluorescence that gradually switched to red after exposure to white light for 1-2 hours. This variant was designated mClavGR1 (monomeric Clavularia-derived green-to-red photoconvertible 1) (Figure 2.2). A complete list of all substitutions present in mClavGR1 and all other variants described in this chapter is provided in Table A1 in Appendix A.

2.2.2 Directed evolution with screening for photoconversion

Starting from mClavGR1 we undertook a total of 11 rounds of evolution, where each round involved library creation by error-prone PCR [182] and StEP shuffling [183] followed by a colony-based screen for those variants with the highest brightness and the highest ratio of red fluorescence after photoconversion to green fluorescence before photoconversion. In each round, approximately 6000 colonies were screened and approximately 5 to 15 of the brightest variants were picked. A mixture of the plasmid DNA from these variants was then used as the template for the subsequent round of library creation and screening. After the 4th round of evolution, we sequenced the clone with the highest fluorescence intensity in both the green (before photoconversion) and red (after photoconversion) channels. This clone, designated mClavGR1.1, was equivalent to mClavGR1 + T6bl, S34R, K36R, K74R, M113I, L166Q, Y173H (Figure 2.2). A genealogy of all mClavGR variants is provided in Figure 2.3A. Directed evolution starting from a template mixture of mClavGR1.1 and other bright variants

was continued for seven more rounds. While we primarily relied on the diversity generated by error-prone PCR and StEP shuffling, we also created libraries in which specific positions that have previously been shown to improve the properties of homologous FPs were randomized (Figure 2.3A and Table A2 in the appendix). These rationally generated libraries were screened in a manner identical to that used for the random libraries.

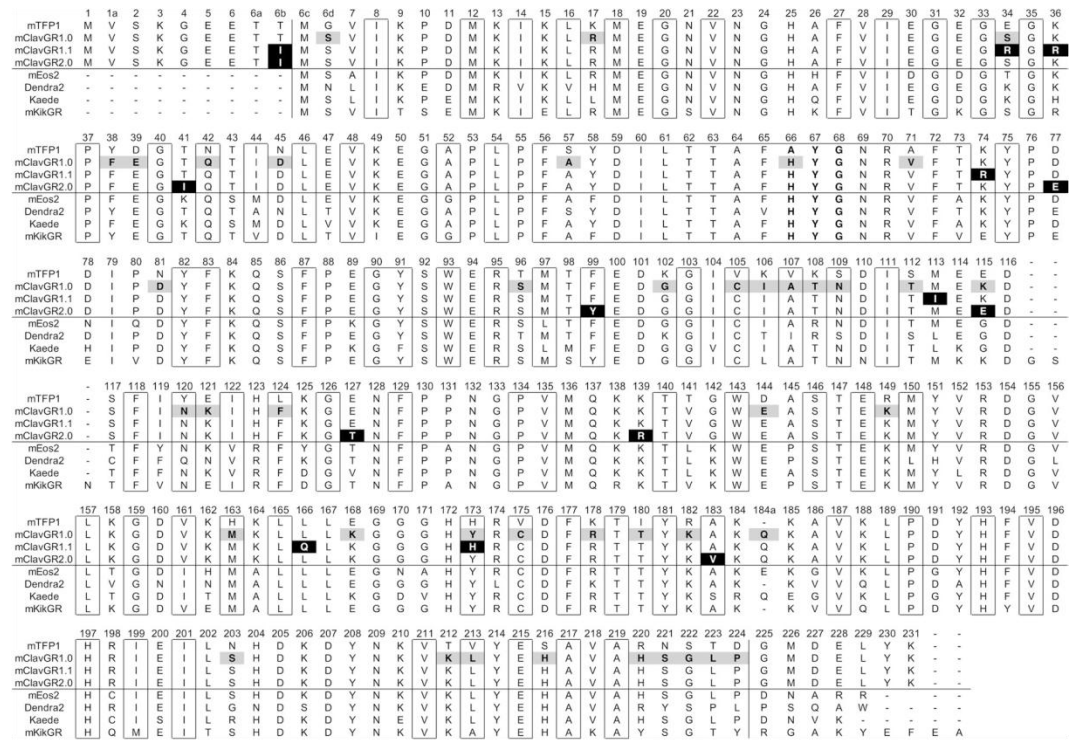


Figure 2.2 Sequence alignment of mClavGR variants and other proteins relevant to this study. Substitutions in mClavGR1, relative to mTFP1, are represented as black text on a gray background. Substitutions in later mClavGR variants that were identified during the directed evolution are shown as white text on a black background. Residues with side chains directed towards the interior of the β -barrel are enclosed in black boxes. The chromophore forming residues are 66-68 in this alignment.

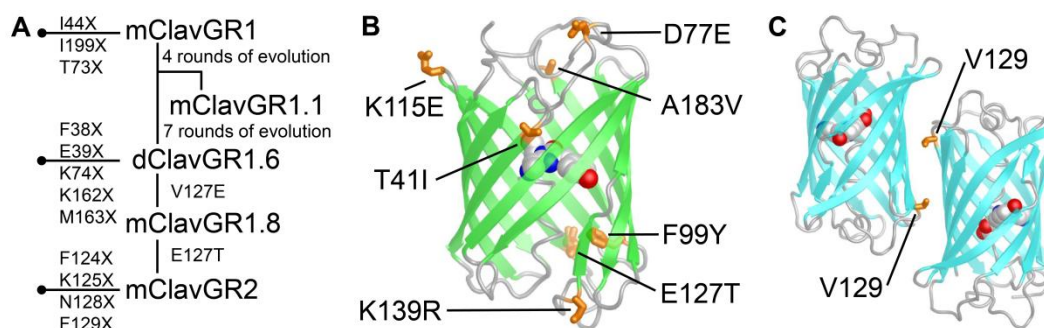


Figure 2.3 Overview of the directed evolution of mClavGR. (A) A genealogy of ClavGR variants. The lines ending in filled circles indicate site-directed libraries that were created and screened, but did not produce improved variants. An ‘X’ represents a selection of amino acids as specified in Table A2 in the appendix. (B) Location of substitutions in mClavGR2 (relative to mClavGR1) that were introduced during the directed evolution process. The X-ray crystal structure of mTFP1 (PDB ID 2HQK) is used here to represent mClavGR [92]. (C) The dimer of amFP486 (from PDB ID 2A48) showing the location of valine 129 which is structurally homologous to residue 127 of mClavGR [92, 94]. As discussed in the text, substitutions at this position can influence the oligomeric state of the protein.

The brightest variant after the 11th round of directed evolution was equivalent to mClavGR1 + T6bl, T41I, D77E, F99Y, K115E, E127V, K139R, A183V (Figure 2.3B). Characterization of the oligomeric state of this variant by gel filtration chromatography revealed that the protein existed as a dimer (data not shown) and so it was designated as dClavGR1.6 where the d prefix indicates a dimer. dClavGR1.6 is about five-fold brighter than mClavGR1 when expressed in *E. coli* and shares only the T6bl substitution with mClavGR1.1. Analysis of the substitutions present in dClavGR1.6, in combination with a structure-based alignment of the X-ray structures of mTFP1 and tetrameric amFP486 [92, 94], led us to suspect that the E127V substitution was most likely responsible for reconstituting the dimer interface (Figure 2.3C). Accordingly, we introduced the single amino acid reversion V127E into dClavGR1.6 by site directed mutagenesis to produce mClavGR1.8. Although mClavGR1.8 was monomeric, it retained only 50% of the fluorescent brightness of dClavGR1.6 when expressed in *E. coli*. Noting that the V123T substitution

had been previously used to convert EosFP into mEos [100], and that position 123 of EosFP aligns with position 127 of mClavGR1.8, we introduce the E127T substitution into mClavGR1.8. The resulting protein, which was designated as mClavGR2 (Figure 2.2 and Figure 2.3B), had regained 60% of the brightness of dClavGR1.6 when expressed in *E. coli* and retained the monomeric character of mClavGR1.8 as determined by gel filtration chromatography (Figure 2.4). Under the same experimental conditions, Dendra2 was also monomeric (Figure 2.4B), but mEos2 existed as a dimer or higher-order oligomer at concentrations of 0.1 mM and greater (Figure 2.4D). Further efforts to improve the protein by screening libraries in which residues of interest were randomized (Figure 2.3A and Table A2), did not result in the identification of further improved variants.

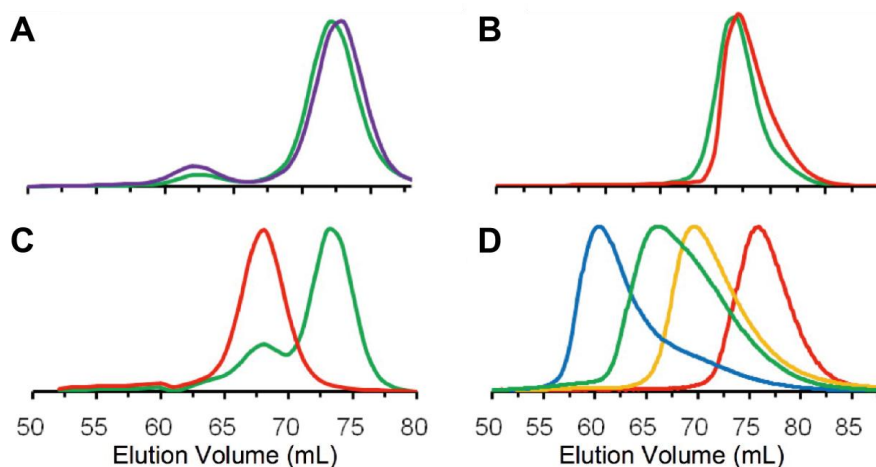


Figure 2.4 Characterization of the oligomeric structure of mClavGR2, mEos2, and Dendra2 by gel filtration chromatography. (A) Coinjection of mCherry and mClavGR2 with detection at both 488 nm (green) and 585 nm (violet). The identity of the species eluting at approximately 63 min is currently unknown, but it was not observed when the sample was first purified with gel filtration as in panel (B). (B) Overlaid elution profiles from separate injections of 1 mM mClavGR2 (green) and 1 mM Dendra2 (red) with detection at 488 nm. (C) Coinjection of mClavGR2 and dTomato with detection at both 488 nm (green) and 555 nm (red). (D) Overlaid elution profiles from separate injections of mEos2 at variant concentrations with detection at 488 nm. The concentrations of the samples are 1 mM (blue), 0.5 mM (green), 0.1 mM (yellow) and 0.01 mM (red). mEos2 existed as a dimer or higher-order oligomer at concentrations of 0.1 mM and greater.

2.2.3 *In vitro* characterization of mClavGR2

As described in the previous section, we had used an empirical screening process in order to produce a protein that had an optimal combination of brightness and photoconversion efficiency. In order to determine what changes in the spectral properties of the protein had resulted in the observed improvements in its properties, we undertook the thorough *in vitro* characterization of mClavGR1 and mClavGR2 (Table 2.1). An overall conclusion from the *in vitro* characterization experiments is that mClavGR2 has spectral properties that are comparable to the other members of the class of green-to-red pcFPs [100]. This result is consistent with the fact that all of the known green-to-red pcFPs, including mClavGR2, have high sequence similarities, particularly when one considers only those residues with side chains directed towards the interior of the β -barrel (discussed below in Section 2.2.4).

Table 2.1 Properties of mClavGR variants and other pcFPs

Protein	State	λ_{ab} (nm)	λ_{em} (nm)	ϵ^a	Φ	Brightness ^b	pK_a	$\Phi_{PC,GRX}/\Phi_{PC,GR1}^c$	$E_{PC,GRX}/E_{PC,GR1}^d$
mClavGR1	green	486	503	16	0.84	14	8.0	1.0	1.0
	red	565	582	21	0.56	12	7.4		
mClavGR2	green	488	504	19	0.77	15	8.0	1.2	1.8
	red	566	583	32	0.53	17	7.3		
mEos2 ^e	green	506	519	56	0.84	47	5.6	ND ^f	ND ^f
	red	573	584	46	0.66	30	6.4		
Dendra2 ^e	green	490	507	45	0.50	23	6.6	ND ^f	ND ^f
	red	553	573	35	0.55	19	6.9		
mKikGR ^e	green	505	515	49	0.69	34	6.6	ND ^f	ND ^f
	red	580	591	28	0.63	18	5.2		

^a Unit of $\text{mM}^{-1} \text{cm}^{-1}$. ^b Product of ϵ and Φ in $\text{mM}^{-1} \text{cm}^{-1}$. For comparison, the brightness of mCherry is $16 \text{ mM}^{-1} \text{cm}^{-1}$ and the brightness of EGFP is $34 \text{ mM}^{-1} \text{cm}^{-1}$ [4]. ^c Photoconversion quantum yield (see Section 2.4.4 for calculation) relative to mClavGR1. ^d Photoconversion contrast (see Section 2.4.4 for calculation) relative to mClavGR1. ^e Previously reported values [100, 127, 180]. ^f Not determined.

Prior to illumination, purified mClavGR2 protein has an excitation maximum at 488 nm and an emission maximum at 504 nm (Figure 2.5A). Illumination with the same 405 nm light source used to photoconvert colonies of *E. coli* causes the formation of a new species with a primary excitation maximum at 566 nm and emission maximum at 583 nm. As represented in Figure 2.1, it is the anionic form of the chromophore that is the primary fluorescent species in both the green and red fluorescent states. Interestingly, the red fluorescent state has a second strong excitation peak at 355 nm. This situation is reminiscent of the excitation spectrum of avGFP, which also exhibits two peaks. In avGFP the higher energy excitation peak is caused by the protonated form absorbing a photon, undergoing excited state proton transfer, and then fluorescing from the anionic excited state [171]. An analogous explanation does not satisfactorily explain the mClavGR2 excitation spectrum, since the protonated form of the mClavGR2 red state chromophore absorbs at 455 nm, 100 nm red-shifted from the 355 nm excitation peak (Figure 2.7C). Adam *et al.* rationalized a similarly positioned peak observed in the excitation spectrum of Dendra2 as being due to a higher energy electronic transition of the red state of the chromophore [93].

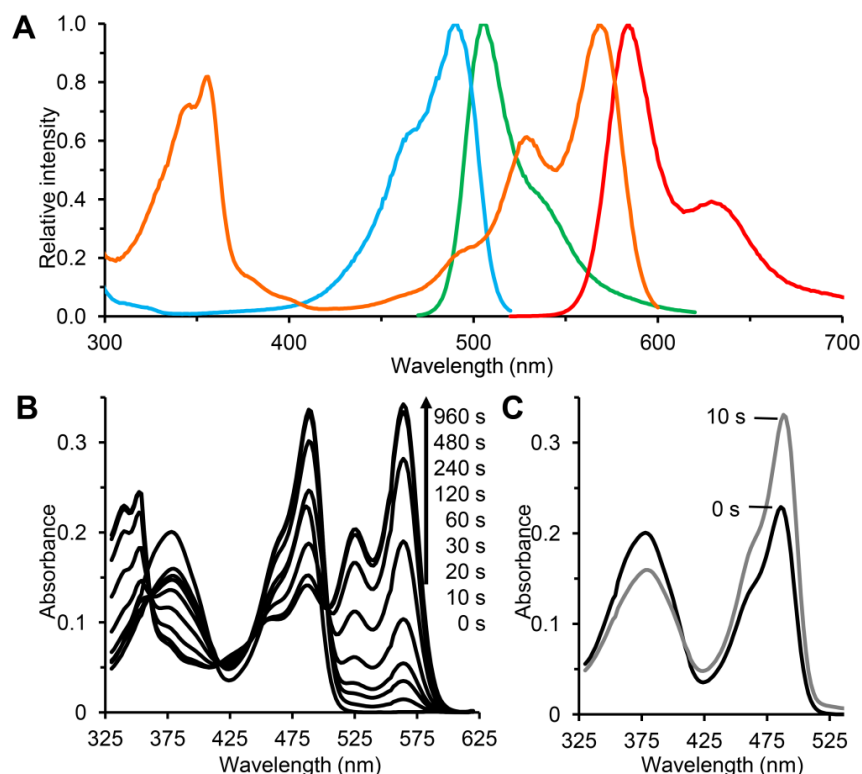


Figure 2.5 Spectral changes in mClavGR2 upon photoconversion. (A) Excitation (cyan and orange) and emission (green and red) fluorescence spectra of the green (pre-photoconversion) and red (post-photoconversion) states of mClavGR2, respectively. (B) Absorbance spectra versus photoconversion time of mClavGR2 in PBS (pH 7.4). The light source used for photoconversion is the 405 nm LED-based illumination chamber used for photoconverting the protein in bacterial colonies. (C) Brief illumination (10 s) with 405 nm light causes a substantial increase in the absorbance of the green fluorescent species.

Prior to illumination, the absorbance spectrum of mClavGR2 is dominated by two distinct peaks (Figure 2.5B); one at 380 nm that is attributable to the neutral phenol form of the chromophore, and one at 488 nm that is attributable to the anionic phenolate form of the chromophore (refer to Figure 2.1). Of these two peaks, only the phenolate form also appears in the excitation spectrum (Figure 2.5A), indicating that only this form is green fluorescent. The neutral phenol form is not green fluorescent but, as previously established by the action spectra of Kaede, EosFP, and Dendra2, is the form that can undergo photoconversion to the red state when excited (refer to Figure 2.1) [93]. Indeed, recording of absorbance

spectra at regular intervals during otherwise continuous illumination with the 405 nm light source revealed a monotonic increase in a 566 nm absorbance peak and a monotonic decrease in the 380 nm peak (Figure 2.5B). The absorbance of the 488 nm peak also undergoes an eventual decrease with continued illumination, illustrating that phenol and phenolate forms of the green state of mClavGR2 are in equilibrium and subject to the law of mass action.

An interesting property of mClavGR2 is that the 405 nm illumination used to photoconvert the neutral phenol form can also temporarily drive the phenol/phenolate equilibrium towards the anionic form. Specifically, we observed that both the absorbance at 488 nm (Figure 2.5C) and the overall fluorescence of the phenolate form increase immediately after illumination with 405 nm light. In the subsequent absence of 405 nm light, the absorbance at 488 nm (and fluorescence at 504 nm) decreases to its original level over several tens of seconds. A similar phenomenon has also been reported to occur with Kaede and Dendra2 [174, 184]. We speculate that this shift in equilibrium could be the result of photo-induced *cis-trans* isomerization of the chromophore in a similar manner to that observed to occur for the green state of the variant of EosFP known as IrisFP [185]. An alternative explanation is that illumination shifts the phenol/phenolate equilibrium towards the phenolate state. However, if this were the case we would expect equilibrium to be reestablished on a timescale of ms as determined for analogous proton transfer processes in avGFP and Kaede, not the tens of seconds that we observe in mClavGR2 [184, 186].

Although both the green and red states of mClavGR2 are approximately 5× brighter than those of mClavGR1 when expressed in *E. coli*, its inherent brightness (i.e. the product of extinction coefficient (ϵ) and quantum yield (Φ)) has improved only by a modest 7% and 42% for the green and red states, respectively (Table 2.1). This indicates that the

improvement of brightness seen in *E. coli* is very likely due to more complete maturation or more efficient folding of the FP. mClavGR2 further benefits from a 20% increase in the quantum yield of photoconversion (Φ_{PC}) relative to mClavGR1. The combination of improved inherent brightness of the red state and the improved Φ_{PC} provides an overall 80% improvement in photoconversion contrast (E_{PC}). E_{PC} is a relative measure of the overall intensity of red fluorescence that can be generated from a fixed concentration of protein molecules.

2.2.4 Structure-based comparison of green-to-red pcFPs

Of the previously reported green-to-red pcFPs, mClavGR2 has the highest amino acid percentage identity with mEos2 (Figure 2.2, Table A3 and Table A4). However, if one considers only those residues with side chains directed toward the interior of the β -barrel, mClavGR2 is most similar to mKikGR with just seven relatively conservative amino acid differences (versus 8 for Kaede, 9 for mEos2, and 13 for Dendra2) (Table A5 and Table A6). Of the green-to-red FPs reported to date, the two that are most similar with respect to the identity of residues directed towards the interior of the β -barrel are mEos2 and Kaede with a difference of only 4. The two that are most divergent are Dendra2 and mKikGR with a difference of 17. As mentioned above, the similar photophysical characteristics of members of this class of FPs is attributed to the fact that they are all very similar with respect to their interior residues and chromophore environments. One of the most distinctive amino acid differences in close proximity to the chromophore is the residue at position 73 (Figure 2.6). In mClavGR2 and Dendra2 (numbered as residue 69) this residue is Thr; in mEos2 and Kaede, it is Ala; and in mKikGR, it is Val (Figure 2.2 and Table A5). It has previously been suggested that hydrogen bonding between the side chain of Thr at this position and the guanidinium group of Arg70 (Arg66 in Dendra2) may hold the guanidinium farther from the chromophore and thus limit charge stabilization [93]. This diminished

degree of charge stabilization in Dendra2 has been used to explain both the blue shifted emission and the higher apparent pK_a of the chromophore.

We attempted to improve the properties of mClavGR1 by introducing substitutions at position 73. When we generated a library of all possible substitutions at position 73 and screened it in colonies to identify the variants that exhibited the highest green fluorescence before photoconversion and the highest red fluorescence after photoconversion, we found that the brightest colonies retained Thr at position 73. We also created the Thr73Ala and Thr73Val variants individually and characterized them. When expressed in *E. coli* the green state of the Thr73Ala variant was approximately equivalent in brightness to the Thr73 progenitor, but the red (photoconverted) state was substantially dimmer. For the Thr73Val variant, both the green and red states were substantially dimmer than the original Thr73 variant.

There are just two residues of mClavGR2 that are directed towards the interior of the β -barrel and have amino acid identities that are not found in any of the other green-to-red pcFPs (Table A5). These residues are Ile44 (Met, Ala, and Val in others) and the Val183 (Ala and Ser in others). The valine at position 183 is the result of an amino acid substitution (Ala183Val) identified through screening of randomly mutated libraries for improved brightness and photoconversion. This is located near one end of the 9th β -strand of the protein and is relatively distant from the chromophore (Figure 2.6). Given its distant location, it is unlikely that Val183 has a direct influence on the properties of the chromophore and is more likely contributing to improved brightness by increasing the folding efficiency of the protein. In contrast, Ile44 is in very close proximity to the chromophore and is likely interacting with the side chain of His66, which is an integral part of the red chromophore (Figure 2.6). We generated a saturation library at Ile44 and again found the best-performing variant remained Ile. We also created the Ile44Met and Ile44Val variants

individually. We found that the Met substitution produced a variant with diminished photoconversion efficiency and the Val substitution produced a variant with dimmer fluorescence in both the green state and the red state.

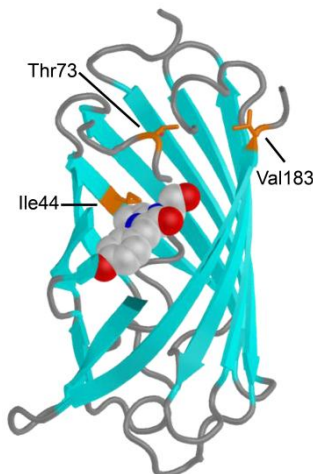


Figure 2.6 Location of positions 44, 73 and 183 that were targeted for semi-saturation mutagenesis. The X-ray crystal structure of mTFP1 (PDB ID 2HQK) is used here to represent mClavGR [92]. Part of the β -barrel has been hidden in order to provide a better view of the positions of interest.

2.2.5 Phenol/phenolate equilibrium and photoconversion

The pH-dependent absorption spectra of the green and red states of mClavGR2 are shown in Figure 2.7. With apparent pK_a s of 8.0 and 7.3, respectively, both states of the chromophore exhibit substantial pH-sensitivity in the physiological pH range (Table 2.1). It is notable that mClavGR2 has the highest pK_a reported for any of the green-to-red pcFPs [100]. Due to our extensive empirical optimization, we expect that the pK_a value for the green species is near to an optimal value with respect to photoconversion efficiency. That is, with this pK_a value, a substantial portion of the protein exists in the neutral phenol form at pH 7.4 and thus the overall rate of photoconversion is increased. A similar relationship between higher pK_a and enhanced photoconversion efficiency has been noted for both KikGR and Dendra2. In the case of KikGR, directed evolution for brighter red fluorescence following photoconversion resulted in a shift of the green state pK_a from an initial value of 4.2 to a final value

of 7.8 [170]. This change dramatically shifted the phenol/phenolate equilibrium towards the neutral form at pH 7.4. In the case of Dendra2 (pK_a 7.1), it was noted that the 20-fold improvement in photoconversion efficiency relative to EosFP (pK_a 5.8) is fully explained by the 20-fold higher concentration of the phenol form at physiological conditions [93]. Presumably, the phenol/phenolate equilibrium in mClavGR2 could be shifted even further to the neutral form in variants with an even higher pK_a . In this case, the intensity of the green fluorescence would have been further diminished. Accordingly, the optimal pK_a of the green state represents a compromise between the desirable but conflicting demands of high photoconversion efficiency and high brightness of the green state.

While the pK_a of the green state of mClavGR2 may be close to optimal, it would be preferable to have a lower pK_a for the red state of mClavGR2. A lower pK_a is beneficial for two reasons. The first is that there are fewer potential artifacts due to local changes in pH. The second reason is that protein is brighter due to a greater fraction of the protein existing in the red fluorescent phenolate form, as opposed to the non-fluorescent phenol form. In mTFP1 (pK_a 4.3), the imidazole side chain of His163 is hydrogen bonded to the phenolate moiety of the chromophore [92, 104]. In mClavGR2, this residue has been substituted with a Met which should be much less effective at stabilizing the anionic form and almost certainly contributes to the dramatically increased pK_a . In an effort to decrease the pK_a of the red state of mClavGR2 we generated the mClavGR2 Met163His variant, even though it seemed likely that this substitution would decrease the pK_a of both the green and red states. Not unexpectedly, the green state of this variant was found to exist entirely in the anionic phenolate form at physiological pH and photoconversion thus proceeded with very poor efficiency.

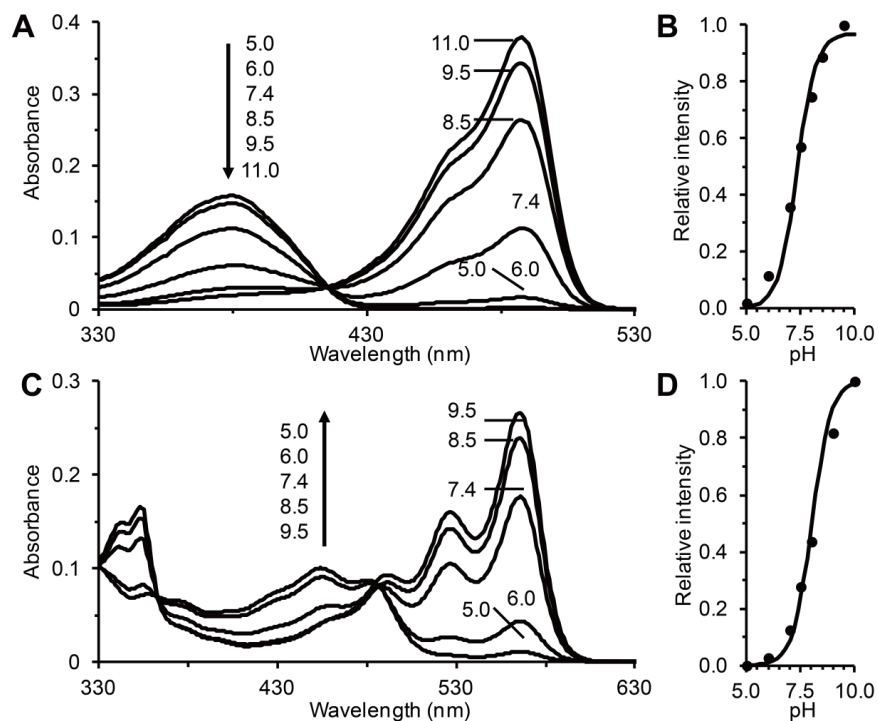


Figure 2.7 Characterization of the pH dependence of the green and red states of mClavGR2. (A) Absorbance spectra of the pre-photoconversion green state of mClavGR2 at pH values ranging from 5 to 11. (B) Plot of relative fluorescent intensity at 490 nm as a function of pH. The line represents the best fit of the data using the Henderson-Hasselbalch equation and is consistent with a pK_a of 8.0. (C) Absorbance spectra of the post-photoconversion red state of mClavGR2 at pH values ranging from 5 to 11. (D) Plot of relative fluorescent intensity at 570 nm as a function of pH and it as in panel (B). The fit is consistent with a pK_a of 7.3.

Given the results of our characterization experiments it is apparent that there are some potential advantages and some potential disadvantages of mClavGR2 relative to other members of the class of green-to-red pcFPs with respect to applications in imaging of dynamic protein localization. Most notable among the advantages of mClavGR2 is its monomeric structure, its highly optimized folding efficiency, and its high photoconversion efficiency due to the high pK_a of the green state. The primary disadvantage is the high pK_a of the red state (Table 2.1 and Figure 2.7) that causes the protein to have approximately half of its maximal fluorescent brightness at physiological pH conditions. It should be noted that, despite having only approximately half of its maximal

brightness, the red state of mClavGR2 retains brightness slightly greater than that of the popular red FP known as mCherry at pH 7.4 (Table 2.1) [103].

2.2.6 Characterization of photostability in live cells

Another advantage of mClavGR2 became apparent while our collaborator Michael Davidson from National Highfield Laboratory (USA) was performing photobleaching experiments with protein expressed in live cells. Specifically, he found that the red state of the protein has improved photostability under confocal illumination conditions, relative to other green-to-red pcFPs (Table 2.2). The time to bleach the red state to 50% of the initial intensity for mClavGR2 (3644 s) represents a substantial improvement over mEos2 (2700 s) and Dendra2 (2420 s) [100]. Oddly, this advantage does not extend to widefield illumination conditions where the time to bleach to 50% of the initial intensity for mClavGR2 (175 s) is somewhat lower than that reported for mEos2 (323 s) and Dendra2 (378 s). It should be noted that the red state of mKikGR exhibits relatively rapid photobleaching under both widefield (21 s) and confocal (530 s) illumination conditions. The time to bleach the green state of mClavGR2 to 50% of its initial intensity using confocal illumination (233 s) is only slightly faster than those of both mEos2 (240 s) and Dendra2 (260 s), but substantially slower than that of mKikGR (80 s). For wide-field photobleaching, mClavGR2 (17 s) is more similar to mKikGR (14 s) than it is to the more photostable mEos2 (42 s) and Dendra2 (45 s) variants.

Table 2.2 Photostability of pcFPs in live cells

Protein	State	Photostability (s) ^a	
		Widefield	Confocal
mClavGR1	green	14	202
	red	64	5002
mClavGR2	green	17	233
	red	175	3644
mEos2	green	42	240
	red	323	2700
Dendra2	green	45	260
	red	378	2420
mKikGR	green	14	80
	red	21	530

^a Time to photobleach from 1,000 to 500 photons s⁻¹ molecule⁻¹ under widefield and confocal illumination, respectively.

2.2.7 Live cell imaging of various mClavGR2 fusion proteins

In an effort to determine if mClavGR2 is potentially useful for imaging of dynamic protein localization, we have assembled mammalian expression plasmids for a variety of mClavGR2 constructs, used these plasmids to transfect cultured mammalian cells, and acquired fluorescence images of the resulting cells both before and after photoconversion. Following expression of mClavGR2 in HeLa cells as an unfused protein, we were able to photoconvert the protein in a localized region and image the subsequent diffusion of the red state protein as it evenly dispersed throughout the cytoplasm over a timescale of approximately 30 s (Figure 2.8). This experiment established that mClavGR2 did not exhibit any anomalous localization when expressed alone in cells and was capable of free and rapid diffusion in the cytoplasm.

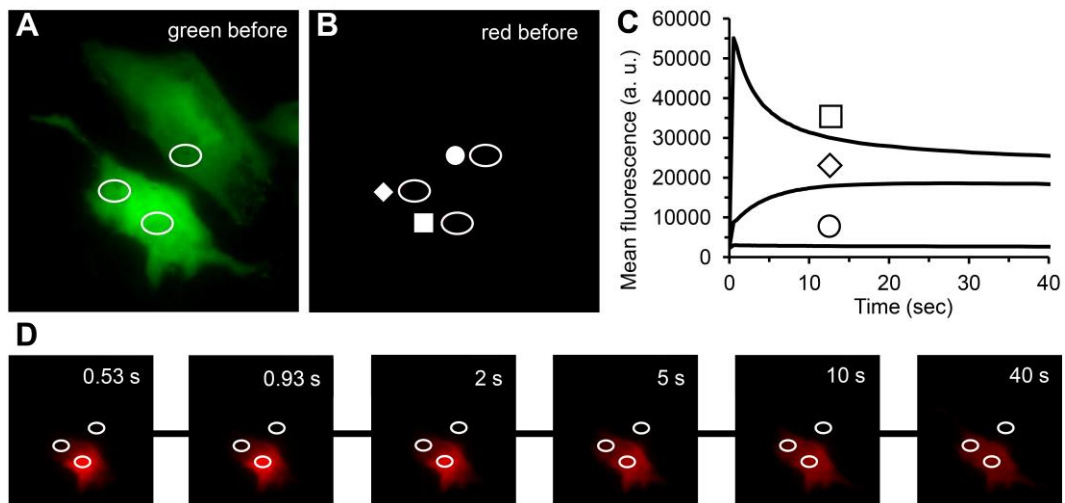


Figure 2.8 Photoconversion of unfused mClavGR2 in live cells. (A) Image of the green fluorescence from two cells expressing mClavGR2 prior to photoconversion. (B) The same cells as in (A) imaged using a filter set for red fluorescence, prior to photoconversion. The three areas of interest are indicated with a circle, diamond, and square. The area indicated with a square corresponds to the location where the cell is illuminated with a 408 nm laser from $t = 0$ to 2 s. (C) Mean red fluorescence intensities in each of the areas of interest marked in panel (B). (D) A series of images showing the free diffusion of photoconverted mClavGR2. Indicated times are elapsed time since the start of the 408 nm laser illumination.

Encouraged by these promising results, our collaborator Michael Davidson helped to construct fusion constructs in order to test whether mClavGR2 could be fused to a different of protein partners and targeting sequences without causing mistargeting or possible misfolding of either protein. In Figure 2.9 we show images of the green state of mClavGR2 in 16 commonly used fusion constructs. These localization patterns for all fusion constructs were judged to be indistinguishable from localization patterns of the analogous *Aequorea* GFP constructs imaged under identical conditions in the same laboratory. The favorable performance of mClavGR2 in these constructs serves as ancillary evidence of its monomeric state. Notably the fine filament structure of the actin cytoskeleton is clearly preserved with the mClavGR2- β -actin fusion expressed in HeLa cells (Figure 2.9A). For the histone 2B (H2B), connexin 43 (Cx43), and Rab5a fusion constructs, localized photoconversion and

subsequent imaging of the dynamic localization of both green and red fluorescence were performed. These experiments established that, when fused to other proteins, mClavGR2 causes minimal perturbation of localization and trafficking dynamics. Furthermore, the green and red states of mClavGR2 have sufficient brightness and photostability for long term imaging experiments. Two-color image sequences for the H2B experiments are provided in Figure 2.10.

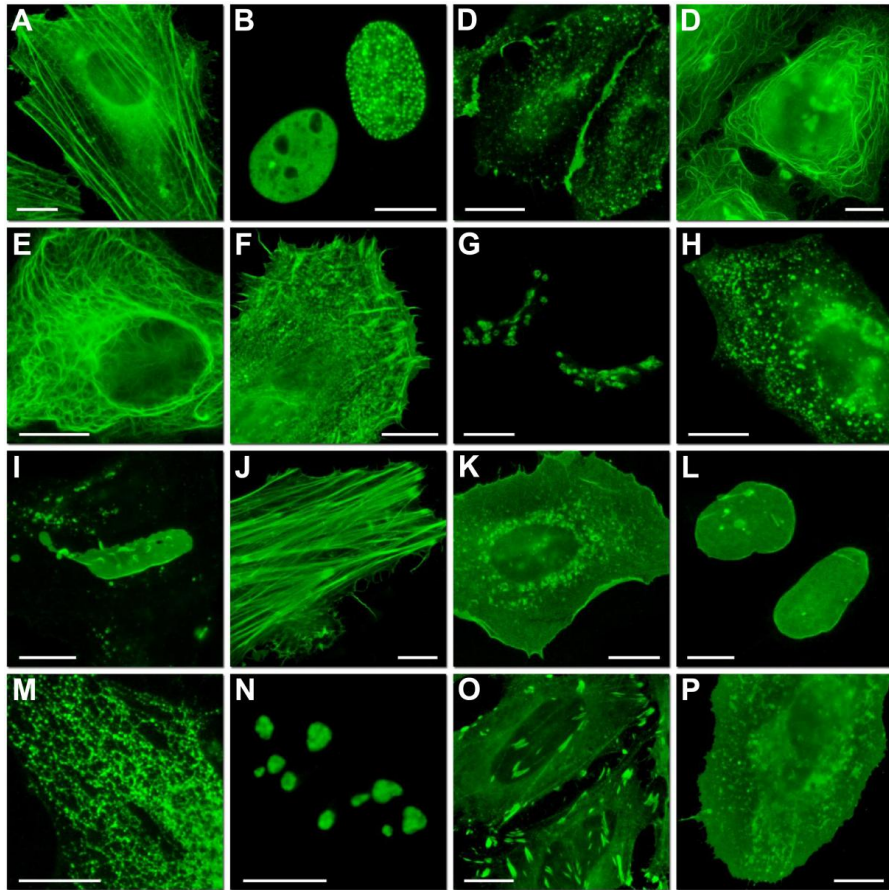


Figure 2.9 Live cell imaging of mClavGR2 protein fusions. The fusion peptide or protein is listed as C-terminal or N-terminal to mClavGR2, followed by the number of amino acids in the linker separating the two domains: (A) mClavGR2-actin-C-7; (B) mClavGR2-CAF1-C-10 (human chromatin assembly factor 1); (C) mClavGR2-VE-cadherin-N-10; (D) mClavGR2-MAP4-C-10 (microtubule binding domain of mouse MAP4); (E) ClavGR2-keratin-N-17 (human cytokeratin 18); (F) mClavGR2- α -actinin-N-19; (G) mClavGR2-GAL-T-N-7 (Golgi complex); (H) mClavGR2-Rab5a-C-7; (I) mClavGR2-Cx43-N-7; (J) mClavGR2-lifeact-N-7 (yeast actin binding protein); (K) mClavGR2-c-src-N-7 (chicken c-src); (L) mClavGR2-lamin

B1-C-10; (M) mClavGR2-ER-N-5 (calreticulin signal sequence N-terminal and KDEL C-terminal); (N) mClavGR2-fibrillarin-C-7; (O) mClavGR2-zyxin-N-6; (P) mClavGR2-c-Ha-Ras (CAAX)-C-5 (farnesylated membrane targeting signal). All fusions were expressed in HeLa cells (CCL-2; ATCC) for imaging. This figure, as well as construction of the mammalian expressing plasmids and imaging are credited to Michael Davidson from National Highfield Laboratory (USA).

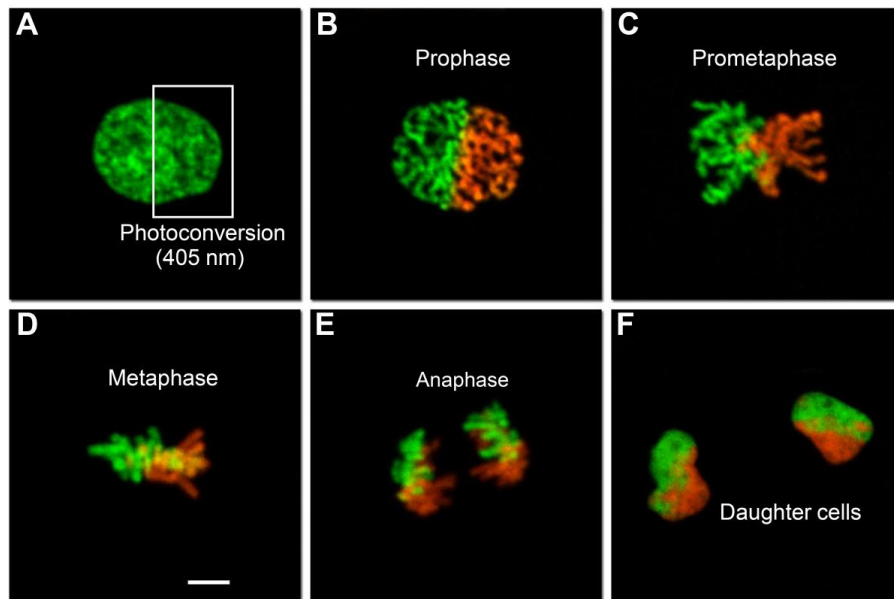


Figure 2.10 Photoconversion and imaging of H2B-mClavGR2. (A) The photoconverted region is indicated with a white rectangle. (B-E) Progression through mitosis which lasts about 2 hours. (F) Daughter cells after mitosis. Imaging credit belongs to Michael Davidson. Scale bar 5 μm .

2.2.8 Imaging protein mobility with mClavGR2

To further test the utility of mClavGR2 and fully exploit the advantages provided by its monomeric structure, we investigated whether mClavGR2 could be used to monitor the dynamics of a transmembrane adhesion protein, intercellular adhesion molecule 1 (ICAM-1). Specifically we asked whether we could quantitatively assess changes in the lateral mobility of ICAM-1 on the surface of a mammalian cell while engaged in a cell-cell contact with a T lymphocyte cell (Jurkat) expressing the ICAM-1 receptors, leukocyte function-associated antigen 1 (LFA-1).

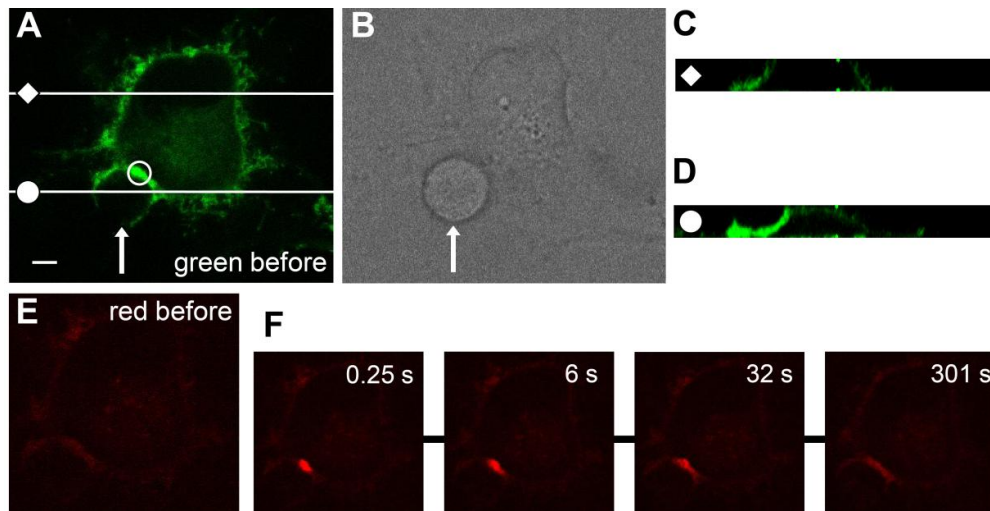


Figure 2.11 Photoconversion of ICAM-1-mClavGR2 on a HeLa cell in contact with a Jurkat cell expressing LFA-1. (A) Green fluorescence image of a HeLa in contact with a Jurkat cell. The arrow indicates the location of the Jurkat cell as shown in panel (B). Brighter fluorescence is observed at the cell-cell contact point due to the interaction of LFA-1 on the surface of the Jurkat with ICAM-1-mClavGR2 on the HeLa cell. The white circle at the cell-cell contact is the area to be photoconverted. Diagonal lines marked with a diamond and a circle correspond to the reconstructed x-z cross-sections represented in panels (C) and (D). Scale bar 5 μ m. (B) Bright field image of the same field of view as panel (A) with focus adjusted to show the Jurkat cell (indicated with white arrow). (C) Cross sectional view of cells shown in panel (A) at the line indicated with a diamond. (D) Cross sectional view of the cell shown in panel (A) at the line indicated with a circle. (E) Red fluorescence image of the same field of view as panel (A) prior to photoconversion. (F) Time-lapse images of the red fluorescence following photoconversion.

Fluorescent images of HeLa cells transiently transfected with a plasmid encoding ICAM-1-mClavGR2 revealed an unambiguous pattern of surface FP localization (Figure 2.11A), consistent with the expected localization of ICAM-1 in the plasma membrane. Incubation of these adherent cells with a suspension of Jurkat cells resulted in the formation of cell-cell contacts where a Jurkat had settled onto the plasma membrane of a HeLa cell. Bright field imaging was used to identify Jurkat cells in apparent physical contact with a transfected HeLa cell (Figure 2.11B). For each apparent contact, the existence of specific ICAM-1/LFA-1 mediated contact was confirmed by locating a bright cluster of fluorescence at the contact point

suggested by the bright field images (Figure 2.11C, D). Localized photoconversion at the point of contact was performed and a time series of red fluorescence images were acquired to trace the diffusion of ICAM-1 initially found at the contact zone (Figure 2.11E). Identical experiments were performed at arbitrary locations on the plasma membrane of transfected HeLa cells in the absence of Jurkat cells. Post-collection analysis was used to extract the fluorescence intensity at the contact point as a function of time (Figure 2.12A).

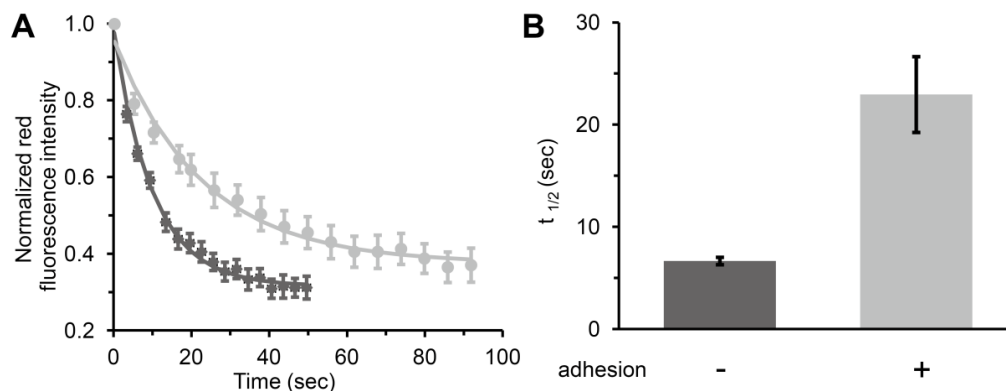


Figure 2.12 Diffusion of ICAM-1-mClavGR2 on the plasma membrane. (A) Red fluorescence intensity at the location of photoconversion versus time after photoconversion for HeLa cells either at the point of adhesion with a Jurkat cell (circles) or at an arbitrary point on the plasma membrane in the absence of Jurkat cells (squares). Data represents the averaged results from 8 and 22 independent measurements, respectively. Lines represent the best fit of the data to $F(t) = A + B\exp(-t/\tau)$. (B) Summary of diffusion times ($t_{1/2}$) as obtained by fitting of the data provided in panel (A). The $t_{1/2}$ values for the two data sets at $p < 0.01$. The “adhesion” label refers to whether or not the site of photoconversion was within an adhesion contact as identified in bright-field images and fluorescence images.

Our analysis of ICAM-1 diffusion revealed that the diffusion half times increased from 6.6 ± 0.4 s for proteins not involved in cell-cell contacts to 22.9 ± 3.7 s for proteins that were involved in cell-cell contacts (Figure 2.12B). This result suggests an almost 4-fold reduction in the diffusion coefficient of ICAM-1 found within a contact zone versus ICAM-1 found elsewhere on the cell surface. Regardless of whether ICAM-1 was in an adhesive contact or not, approximately 25% ($24 \pm 4\%$ or $26 \pm 2\%$,

respectively) of the protein was effectively immobile over the 100 s timescale of the experiment (Figure 2.12B). These results are consistent with previous reports that ICAM-1 mobility is reduced by crosslinking of the receptor, although here crosslinking is induced by the presentation of receptors as part of the lymphocyte cell surface [187]. These results unambiguously demonstrate that photoconversion of mClavGR2-containing constructs will be a powerful tool for dissecting the mechanisms that mediate cellular adhesion by allowing long-term tracing of receptors and lateral diffusion measurements.

2.2.9 Engineering and characterization of mMaple

Though mClavGR2 is monomeric at high concentrations (Figure 2.4), we were concerned that it might have a weak tendency to dimerize, since residues 220 – 224 (HSGLP) are identical to the ones that form part of the dimer interface in related Anthozoa FPs [94]. Thus, we replaced residues 220 – 224 with the corresponding residues (RNSTD) from the progenitor protein mTFP1 [92] (Figure 2.13).

Starting from this modified mClavGR2 variant, we undertook 4 rounds of protein optimization by creating successive libraries of ~5 thousand genetic variants and screening these libraries for efficient photoconversion. A shorter photoconversion time were used in the screening and *E. coli* colonies with high brightness and an improved ratio of red fluorescence (post-photoconversion) to green fluorescence (pre-photoconversion) were considered “winners” of a given round of screening. In the final round of optimization, the winners of all previous rounds were genetically shuffled [183]. Screening of this final library led to the discovery of mMaple, a variant that retains many of the key traits of mClavGR2 (Figure 2.14 and Table 2.3), yet provides an improved ratio of red-to-green during photoconversion (Figure 2.15B). mMaple is equivalent to mClavGR2 with the HSGLP to RNSTD replacement and A145V/G171S/G225S.

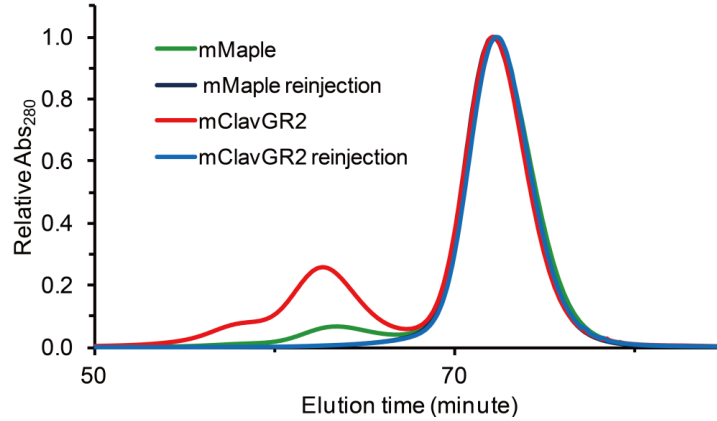


Figure 2.13 Gel filtration chromatography of mMaple and mClavGR2. Both mClavGR2 (injection concentration of 0.5 mM) and mMaple (injection concentration of 0.5 mM) purified from *E. coli* by Ni²⁺-NTA affinity chromatography show an additional peak at 63 min (also see Figure 2.4A). This peak is diminished in size for mMaple relative to mClavGR2 (6.7% vs. 25.8% of monomer peak area). While this peak does elute at a time consistent with the dimer species, it was not observed following reinjection of the collected and concentrated monomeric peak of either mClavGR2 (injection concentration of 0.34 mM) or mMaple (injection concentration of 0.39 mM). The fact that the species eluting at 63 min was not observed in the reinjection suggests that it is not the typical non-covalent dimer species expected for weakly dimerizing fluorescent proteins. While the nature of this species remains unclear, it is apparent that the tendency of it to form is reduced in mMaple.

Table 2.3 Properties of mMaple and related variants

Protein	State	λ_{ex} (nm)	λ_{em} (nm)	ϵ^{a}	Φ	Brightness ^b	pK_{a}
mMaple	green	489	505	15 (26, 59)	0.74	11	8.2
	red	566	583	30	0.56	17	7.3
mClavGR2	green	488	504	19 (25, 60)	0.77	15	8.0
	red	566	583	32	0.54	17	7.3
mEos2	green	506	519	78 (32, 95)	0.43	34	5.6 ^c
	red	573	584	39	0.35	14	6.4 ^c

^a Extinction coefficient ($\text{mM}^{-1} \text{cm}^{-1}$) at peak absorbance wavelength in PBS (pH 7.4). Value in parentheses was determined at pH 4 and pH 10, respectively. ^b Product of ϵ and Φ in $\text{mM}^{-1} \text{cm}^{-1}$. For comparison, the brightness of EGFP and mCherry are $34 \text{ mM}^{-1} \text{cm}^{-1}$ and $16 \text{ mM}^{-1} \text{cm}^{-1}$, respectively [4]. ^c Data from McKinney *et al.* [100].

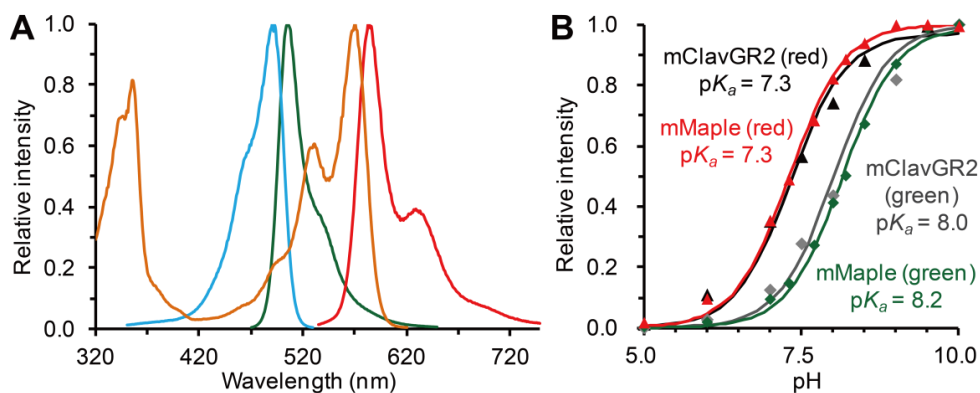


Figure 2.14 Spectral characterization and pH titration of mMaple. (A) Excitation (cyan and orange) and emission (green and red) spectra of the pre-photoconversion green state and the post-photoconversion red state of mMaple. They are indistinguishable from the fluorescence spectra of mClavGR2. (B) pH titration of mMaple and mClavGR2. The green state of mMaple has a pK_a of 8.2 (green diamond), slightly higher than that of mClavGR2 (8.0, grey diamond). However, the red state of both FPs has an identical pK_a of 7.3 (black triangle for mClavGR2, red triangle for mMaple).

In vitro characterization revealed that the primary difference between mMaple and the earlier mClavGR2 variant is a shift in the ground state equilibrium of the green state chromophore from the phenolate form (absorbance $\lambda_{\max} = 489$ nm) towards the phenol form (absorbance $\lambda_{\max} = 380$ nm) (Figure 2.15). This shift is attributed to an increase in the apparent pK_a of the green state chromophore from 8.0 to 8.2 (Figure 2.14B and Table 2.3). The increased ratio of phenol to phenolate form in the green state explains the improved photoconversion contrast of mMaple (Figure 2.15B), since it is the phenol form that undergoes the green-to-red photoconversion and it is the phenolate form that is green fluorescent. The post-conversion red state retains the same pK_a as mClavGR2 (7.3), so the population of the red fluorescent phenolate form remains unchanged (Figure 2.14B and Table 2.3). We speculate that the A145V mutation is primarily responsible for the shift of the green state pK_a , since position 145 is located immediately adjacent to the tyrosine-derived phenolate moiety of the chromophore. Although it does not directly interact

with the chromophore, the bulkier side chain of valine may stabilize the protonated state by decreasing the solvent accessibility of the chromophore. Notably, position 145 is occupied by proline in all other pcFPs except Kaede [75], which has alanine at this position. The effect of the additional mutations (G171S, G225S) is unclear, as they are relatively remote from the chromophore.

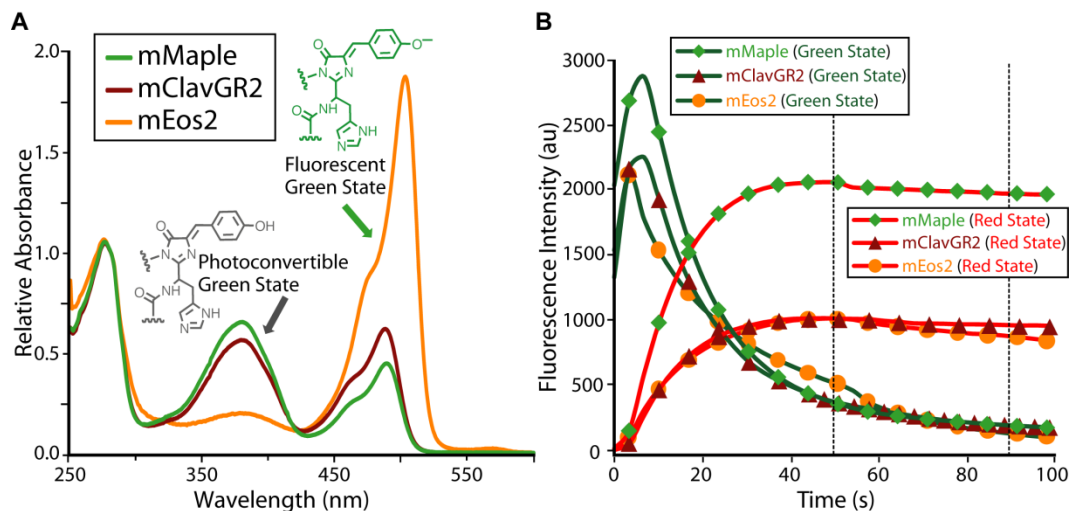


Figure 2.15 Characterization of mClavGR2 and its improved variants. (A) Absorbance spectra of mMaple (green line), mClavGR2 (red line) and mEos2 (orange line). Absorbance spectra are normalized to absorbance at 280 nm. (B) Photoconversion from the green state (green lines) to the red state (red lines) of mMaple (diamonds), mClavGR2 (triangles), and mEos2 (circles) for pcFP-H2B fusions expressed in HeLa S3 cells. The red-to-green contrast has been calculated at 47 s and 91 s (indicated with vertical dotted lines). At 47 s the contrast is 5.3 for mMaple, 2.1 for mClavGR2, and 1.8 for mEos2. At 91 s the contrast is 10.8 for mMaple, 4.7 for mClavGR2, and 6.9 for mEos2. Photoconversion experiment in (B) was performed by Michael Davidson. Figure was prepared in collaboration with Ann McEvoy.

2.3 Conclusions

We have developed and characterized a new green-to-red pcFP variant designated as mClavGR2 and its derivative mMaple. Unlike other members of this class of FP, mClavGR2 was engineered from a well-characterized and monomeric FP progenitor. Our results demonstrate that both mClavGR2 and mMaple has favorable spectral properties, retain the

monomeric structure of its progenitor mTFP1 and, in the constructs tested to date, does not interfere with the correct localization of a genetically fused protein partner. We anticipate that these new additions to the toolbox of engineered FPs will be of great utility in imaging of fast protein dynamics in live cells. Experiments to determine whether the advantages of mClavGR2 and mMaple translate to improved performance in super-resolution imaging applications [177] will be discussed in next chapter.

2.4 Materials and methods

2.4.1 General methods and materials

All synthetic DNA oligonucleotides for cloning and library construction were purchased from Integrated DNA Technologies (Coralville, IA). *Taq* (New England Biolabs) and *Pfu* polymerases (Fermentas) were used for error-prone PCR or regular PCR, respectively. PCR products and products of restriction digests were purified using QIA gel extraction kit (Qiagen) or GeneJET gel extraction kit (Fermentas) according to the manufacturer's protocols. Restriction enzymes were purchased from New England Biolabs or Fermentas. The cDNA sequences were confirmed by dye terminator cycle sequencing using the DYEnamic ET kit (Amersham Biosciences) or BigDye Terminator v3.1 Cycle Sequencing Kit (Applied Biosystems). Sequencing reactions were analyzed at the University of Alberta Molecular Biology Service Unit or the Florida State University Department of Biological Science DNA Sequencing Facility.

2.4.2 Plasmid library creation and site-directed mutagenesis

Randomly mutated libraries were created by error-prone PCR, where dNTPs mix with a deficient component were used and $MnCl_2$ added to the reaction to further decrease the fidelity of *Taq* polymerase [182]. Saturation/semi-saturation mutagenesis at specific residues was performed using over-extension PCR [188]. In the first step of the over-

extension PCR, two PCR were run: one using a 5' primer for the full-length gene and a 3' primer containing the desired mutations at the target site; the other PCR using a 5' primer which shares an overhand with the 3' primer in the first reaction and anneals the target site and a 3' primer for the full-length gene. In the second step, a mixture of the PCR product from the first step were use as template and 5' and 3' primer for the full-length gene were use to amplify the full-length gene incorporating the desired mutations. Staggered extension process (StEP) PCR was performed as previously described [183]. Specifically, a total of 0.1 pmol pBAD/His B plasmids containing DNA sequence of 5-15 mClavGR variants selected from randomly mutated libraries were used as the template mix. The mixture was subjected to 100 cycles (94 °C for 45 s then 55 °C for 5 s), followed by an extra extension step at 72 °C for 5 min. Regardless of library assembly method (i.e. error-prone PCR, saturation mutagenesis, or StEP PCR), gene libraries were inserted into pBAD/His B between the *XhoI* and *EcoRI* sites as described above. Electrocompetent *Escherichia coli* (*E. coli*) strain DH10B (Invitrogen) was transformed and plated on LB (Luria–Bertani)/agar plates supplemented with ampicillin (0.1 mg/ml) and L-arabinose (0.2%). Plates were incubated for 14 h at 37 °C prior to screening.

2.4.3 Plasmid library screening

E. coli colonies expressing the mClavGR libraries were grown on 10 cm Petri dishes. In order to screen libraries for variants that were brighter and exhibited more efficient photoconversion, we developed a method for photoconverting FPs when expressed in colonies of *E. coli*. Our method involves a three-step process. In the first step, a custom imaging system equipped with filter sets (Chroma) for 470/40 nm excitation with 510/20 nm emission (i.e. for green fluorescence) and 560/40 nm excitation with 630/60 nm emission (i.e. for red fluorescence) is used to acquire both green and red fluorescence images of the Petri dish. In the second step,

the dish is placed in a “photoconversion chamber” where it is evenly illuminated with the light from six 9 × 11 arrays of 405 nm light emitting diodes (LEDs) (OptoDiode Corporation, Newbury Park, CA). Following 10 to 20 min of illumination, the dish is imaged again using the same filters that were used in the first step. The digital images are loaded into Image Pro plus software (Media Cybernetics, Bethesda, MD) where they are aligned and processed using a custom macro. The output of the macro is a list of the mean intensity for each colony in each of the four images. This data is imported into Microsoft Excel where it is plotted as *green fluorescence before photoconversion vs. red fluorescence after photoconversion*. During the directed evolution process of mClavGR2, colonies that exhibited the highest intensities in both channels were picked and used as the templates for the following round of library creation. For directed evolution towards mMaple, colonies with an improved ratio of red fluorescence to green fluorescence were picked.

2.4.4 Protein purification and characterization

For production of protein, *E. coli* strain DH10B was transformed with the pBAD/His B expression vector containing the FP gene of interest. A single colony was used to inoculate a 4 mL LB culture supplemented with ampicillin and was grown overnight at 37°C with shaking at 225 rpm. The overnight culture was diluted 1:1000 into LB medium supplemented with ampicillin and grown until the culture has an OD of 0.6. At that time, L-arabinose (final concentration 0.02%) was added to induce the protein expression. After another 12-24 hours of growing, the cells were harvested by centrifugation and lysed by cell disrupter. Proteins were purified by Ni-NTA chromatography (Amersham) and the buffer was changed into phosphate buffered saline (PBS, pH 7.4).

Molar extinction coefficients (ϵ) of the green states were measured by the alkali denaturation method [103, 189]. Briefly, the protein was diluted into regular buffer or 0.01 M NaOH and the absorbance spectra recorded

under both conditions. The ε was calculated as: $\varepsilon = (Abs_{sample}/Abs_{den,445}) \times \varepsilon_{den,445}$, where $\varepsilon_{den,445}$ is 44,000 M⁻¹ cm⁻¹. To determine ε for the red state, the green state protein absorbance was recorded and the sample was then photoconverted until the absorbance peak of the red species reached a plateau. The ε of the green state was used as reference to calculate that of the red states. Fluorescence quantum yields (Φ) were determined using fluorescein in 10 mM NaOH ($\Phi = 0.95$) [190] and Rhodamine 6G in ethanol ($\Phi = 0.94$) [191] as standards.

For determination of the pH dependence, purified protein in PBS was diluted 1:50 into a series of pH-adjusted citrate saline (pH \leq 8) or sodium phosphate (pH $>$ 8) buffers in a 96-well black clear bottom plate (Corning). Fluorescence was measured using a Safire2 plate reader (Tecan).

To determine the relative quantum yields for photoconversion, each purified variant's absorbance at 487 nm was measured to determine the concentration. The proteins were then converted into the red state by the same illumination system used for screening for one minute. The absorbance at 565 nm of the photoconverted proteins was measured to determine the concentration of the photoconverted red species. The proteins were then diluted 10-fold into PBS and their red fluorescence was recorded. The photoconversion quantum yield relative to mClavGR1 ($\Phi_{PC,GRX}/\Phi_{PC,GR1}$) was calculated as

$$\frac{\Phi_{PC,GRX}}{\Phi_{PC,GR1}} = \frac{c_{red,GRX} / c_{green,GRX}}{c_{red,GR1} / c_{green,GR1}} = \frac{A_{565,GRX} / \varepsilon_{565,GRX}}{A_{487,GRX} / \varepsilon_{487,GRX}} * \frac{A_{487,GR1} / \varepsilon_{487,GR1}}{A_{565,GR1} / \varepsilon_{565,GR1}}$$

, where A_{487} , A_{565} , ε_{487} , and ε_{565} , are the absorbance and extinction coefficients at 487 nm (before photoconversion) and 565 nm (after photoconversion), respectively. "GR1" refers to mClavGR1, while GRX refers to one of the other variants. This relative photoconversion quantum yield gives the frequency of generation of the red species of the evolved variants relative to the parent mClavGR1 when illuminated by the screening system. A relative photoconversion effectiveness ($E_{PC,GRX}/E_{PC,GR1}$) was calculated

using $\frac{E_{PC,GRX}}{E_{PC,GR1}} = \frac{c_{GR1}}{c_{GRX}} * \frac{I_{GRX}}{I_{GR1}}$, again with mClavGR1 as reference. In the equation, c is the concentration of the purified protein solution and I is the red fluorescence intensity following photoconversion. This relative photoconversion efficiency can differ from the photoconversion yield due to additional factors such as the ratio of neutral form versus anionic form at the beginning and the brightness of the red state. All absorption measurement was acquired on a DU-800 UV-visible spectrophotometer (Beckman). All fluorescence spectra were recorded on a QuantaMaster spectrofluorimeter (Photon Technology International) and have been corrected for the detector response.

The oligomeric structure of mClavGR variants was determined by gel filtration chromatography with a HiLoad 16/60 Superdex 75 pg gel filtration column on an AKTAbasic liquid chromatography system (GE Healthcare) or a AKTA Prime Plus system (GE Healthcare). Samples of the dimeric dTomato and the monomeric mCherry proteins [103] were expressed and purified in the same way as mClavGR variants and used as size standards. Purified mClavGR variants were mixed with dTomato or mCherry and the resulting elution profiles were monitored at 488 nm for mClavGR variants, 555 nm for dTomato, and 585 nm for mCherry.

2.4.5 Live cell imaging

Construction of mammalian expression plasmids. To create the mClavGR2 vector for cytoplasmic expression, the mClavGR2 gene was inserted between the *XhoI* and *EcoRI* restriction sites of a modified pcDNA3 vector (Invitrogen) in which the multiple cloning site had been modified to contain *XhoI* and *EcoRI* restriction sites in the same reading frame as the identical sites in pBAD/His B. The full length ICAM-1 (Addgene plasmid 8632) [192] and mClavGR2 cDNA were incorporated into the modified pcDNA3 vector mentioned above, with a 24 amino acids

linker between them. The linker's sequence is KALSAAAGGGGSGGGGSGGGGSEF [193]. Plasmid DNA for transfection was prepared by Qiagen Plasmid Midi Kit (Qiagen) according to the manufacturer's protocol.

Mammalian cells maintenance and transient transfection. HeLa cells (CCL2 line; ATCC) were maintained in Dulbecco's Modified Eagle Medium (DMEM) (Invitrogen) supplemented with 10% heat-inactivated fetal calf serum (FCS) (Invitrogen) and 2 mM GlutaMax (Invitrogen) at 37 °C and 5% CO₂. Jurkat cells (clone E6.1) were maintained in Roswell Park Memorial Institute medium-1640 (RPMI-1640) (Invitrogen) supplemented with 10% FCS. Cells for transfection and imaging were cultured on 35 mm glass-bottom culture dishes. Lipofectamine 2000 (Invitrogen) was used for transfection following the manufacture's protocol. The medium was changed into HEPES-buffered Hank's Buffered Salt Solution (HHBSS) (Invitrogen) before imaging. Imaging was done 24-36 h after transfection.

Live cell microscopy. Fluorescence microscopy of mClavGR2 diffusion in the cytoplasm was performed using an inverted Nikon Eclipse Ti microscope equipped with an QuantEM 512SC electron-multiplying CCD (EMCCD) (Photometrics), a 75 W Xenon lamp for epi-fluorescence illumination, a 63× oil-immersion objective (Nikon), and a 408 nm photoactivation laser with a power of >500 mW (Melles Griot). The photoactivation laser delivers a point illumination with an approximate diameter of 40 pixels at the center of the field of view. The following filters were used: 465-495 nm (excitation), 515-555 nm (emission), and 505 nm (dichroic) for the green channel; 528-553 nm (excitation), 590-650 nm (emission), and 565 nm (dichroic) for the red channel. Image procession was done using ImageJ [194].

Imaging of ICAM-1-mClavGR2 diffusion. Jurkat cells were pretreated with phorbol 12-myristate 13-acetate (PMA) at a concentration of 50 ng/mL for 30 min at 37 °C. The stimulated Jurkat cells were centrifuged

and resuspended in DMEM/10%FCS/GlutaMAX. Approximately 10^6 Jurkat cells were added to a 35 mm glass-bottomed culture dish containing HeLa cells previously transfected with the ICAM-1-mClavGR2 plasmid. The cell mixture was incubated at 37 °C for another 30 min to allow adhesion and the solution was exchanged to remove Jurkat cells remaining in suspension. The imaging was performed on a PerkinElmer Ultraview ERS system (PerkinElmer Life Sciences Inc., MA, USA) equipped with a Yokogawa (Yokogawa Corp. Japan) Nipkow spinning disk confocal system, a 60× oil-immersed objective (Carl Zeiss), an EMCCD (Hamamatsu) and a 20 mW 405 nm laser (Melles Griot). Temperature was maintained at 37 °C during the acquisition. An exposure of approximately 40 ms at 15% laser power (405 nm) was used to photoconvert mClavGR2 within the defined area of interest. Four images of each of the green and red channels were acquired before photoconversion, followed by 10 seconds of acquisition at maximum speed, followed by 2 min at 3 seconds per frame for 2 min. The average intensity of the four images acquired before photoconversion was subtracted from the background corrected red fluorescence intensity, which was then normalized to the first image after photoconversion. The fluorescence as a function of time ($F(t)$) was fit with $F(t) = A + B * \exp(-t / \tau)$, where A and B correspond to the fractional contributions of immobile and mobile protein, respectively; and τ is the time constant for the decay of fluorescence intensity. The diffusion half time ($t_{1/2}$) equals $0.69 \times \tau$.

Chapter 3 Advanced application and characterization of mClavGR2 and mMaple³

³ The research described in this chapter is a close collaboration of several groups and has been published on *PLoS One* [169]. Experiments performed by me include the *in vitro* biochemical characterization (Figure 3.4, 3.5, and 3.7), the mammalian expression studies (Figure 3.10), and data analysis for the photostability characterization (Figure 3.8 and 3.10).

3.1 Introduction

As mentioned in the previous chapter, an exciting application of photoconvertible FPs is that they can serve as probes for single molecule localization-based super-resolution microscopy techniques such as photoactivated localization microscopy (PALM) [195], stochastic optical reconstruction microscopy (STORM) [196], and fluorescence-PALM (f-PALM) [197]. These techniques will collectively be referred to as (f-)PALM/STORM hereafter. The lateral resolution of conventional fluorescence microscopy is limited to ~ 200 nm for the shortest wavelength of visible light (~ 400 nm) as stated in Abbe's law [198]. When imaged with a microscope, the diffraction of light causes an object with dimensions smaller than the resolution (i.e. a point source) to appear on the image plane as an Airy pattern, a bright disk (known as Airy disk) surrounded by a series of weakening concentric bright rings (Figure 3.1A). The intensity profile of the Airy pattern of a point source is known as the imaging system's point spread function (PSF). The radius of the Airy disk (r_{Airy}) depends on the wavelength of the light and the numerical aperture of the objective. Two point source objects located close together in the sample can only be individually distinguished (i.e. resolved) in the image when the distance between their centers is greater than r_{Airy} (Figure 3.1B). In other words, when two adjacent objects have overlapping Airy disks, they can no longer be resolved. This fundamental limit of the ability of light microscopy to resolve two adjacent objects is commonly referred to as Abbe's resolution limit or the diffraction barrier.

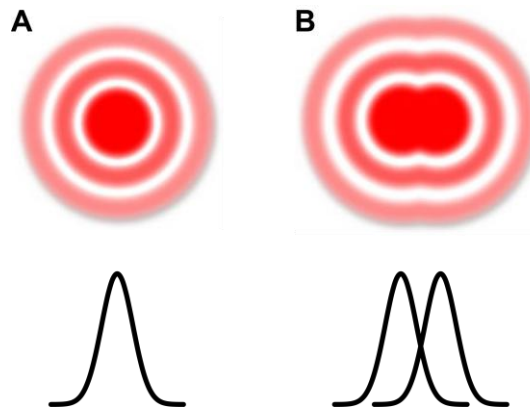


Figure 3.1 Lateral resolution of optical microscope. (A) Representative Airy diffraction image of a sharp-point-like object (top) and the one-dimensional intensity profile (bottom) of its Airy disk. (B) The diffraction-limited image pattern of two point objects that are separated by a distance of r_{Airy} and justified to be resolvable by Rayleigh criterion. Image of Airy diffraction pattern adapted from a microscopy tutorial on the Zeiss website (<http://zeiss-campus.magnet.fsu.edu>).

More than one century after the introduction of the concept of Abbe's resolution limit, researchers have finally found several different approaches to overcome this diffraction barrier in optical microscopy. Microscopy techniques such as (f-)PALM/STORM that circumvent the classical diffraction barrier are thereby referred to as super-resolution microscopy. (f-)PALM/STORM take advantage of photoactivatable and photoconvertible fluorescent probes. By stochastically turning on a subset of the probes and imaging them, the localization of each molecule can be pinned down via computationally-assisted identification of the centres of the PSFs. Hundreds or thousands of photons from each fluorescent probe can be collected before it eventually is destroyed by irreversible photobleaching. This "mapping" process is repeated until all probes have been activated, imaged and photobleached. Usually 100-1000 frames of sparsely distributed single molecules are recorded, resolved and assembled in to a super-resolution image with 20-30 nm lateral resolution [177]. Therefore, to some extent, the improvement in the spatial resolution comes at the expense of temporal resolution as it typically takes several minutes to acquire a full (f-)PALM/STORM image [177]. Recent technical

improvements allow effective acquisition of a complete frame at the timescale of seconds [199, 200].

Other superresolution microscopy techniques such as stimulated emission depletion (STED) [201, 202] and structured illumination microscopy (SIM) [203, 204] attempt to surpass Abbe's limit by playing "tricks" with the illuminating light. Similar to emission, the pattern of excitation light is also limited by the diffraction rule. In STED, a second red-shifted laser, typically with a donut-shaped beam that surrounds the beam of the excitation laser, is used to quench the fluorescence at the periphery of the excitation beam. This is achieved by driving the excited fluorophores at the periphery of the excitation beam into the ground state via stimulated emission. This reduces the effective size of excitation light beam, as only the molecules at the innermost of the excitation beam can effectively emit fluorescence. Using GFP as the fluorescent tag, STED imaging with resolution of ~ 70 nm has been demonstrated [205]. SIM overcomes Abbe's law by using periodically patterned illumination that interacts with the structured sample to generate an interference image. Higher-order frequency information in the sample that was previously not resolvable can then be extracted from the interference image [203]. The observable spatial frequency of the sample in SIM is increased up to a factor of two as the frequency of the excitation pattern itself is limited by diffraction. Since structures parallel to the illumination patterns do not benefited from shifted illumination, a series of images at different orientation are required to reconstitute the full image.

STED and SIM provide the advantage of relatively high speed image acquisition, while (f-)PALM/STORM generally provided better spatial resolution [177, 206]. Since each super-resolution microscopy technique has its own advantages and limitations, it would be beneficial if the same sample can be imaged with different modalities. However, different probe properties are desired in different techniques--while fluorescent probes with high brightness and photostability are desired in STED and SIM,

photoactivable or photoconvertible probes must be used in (f-)PALM/STORM.

Our collaborator Dr. Ann McEvoy from the lab of Jan Liphardt at the University of California, Berkeley, has successfully applied mMaple in super-resolution imaging modalities including (f-)PALM/STORM and SIM and compared its performance with that of other pcFPs. These results are briefly summarized here, but further details are available in our accepted manuscript by *PLoS One* [169]. For effective comparison, the relatively well-characterized chemotaxis protein CheW was used as a model. The expression level of FP-CheW fusion proteins in *E. coli* is controlled by an L-arabinose inducible promoter and the functionality of CheW-FP fusions is tested by swarming assay. Following experimental verification that each pcFP-CheW fusion recovers the chemotaxis ability in a *cheW* knockout strain, the subcellular CheW distribution was imaged by localizing hundreds of single proteins per cell with PALM (Figure 3.2A-C). These imaging results showed that mMaple-CheW gave a substantially higher number of single molecule localizations (930 ± 540) than either mClavGR2 (400 ± 180) or mEos2 (270 ± 110) in the *cheW* knock-out *E. coli* strain. The number of observable localization in the mMaple-CheW construct is much closer to the native expression level of CheW protein (~ 2400 considering that only the bottom $\sim 40\%$ of the cell is covered by the depth of field of objective) [207]. Noticeably, the mean number of photons emitted in the red state, which is a measurement of localization precision [208], is similar in all pcFPs (Figure 3.2D). To ensure that the higher number of localizations observed for mMaple-CheW is not fusion specific, *E. coli* expressing pcFPs with no fusions were imaged and counted in a similar way. In this case, approximately 10 \times the number of localizations per cell was observed for mMaple (3500 ± 1600 localizations per cell) relative to mEos2 (210 ± 90 localizations) (Figure B1 in Appendix B).

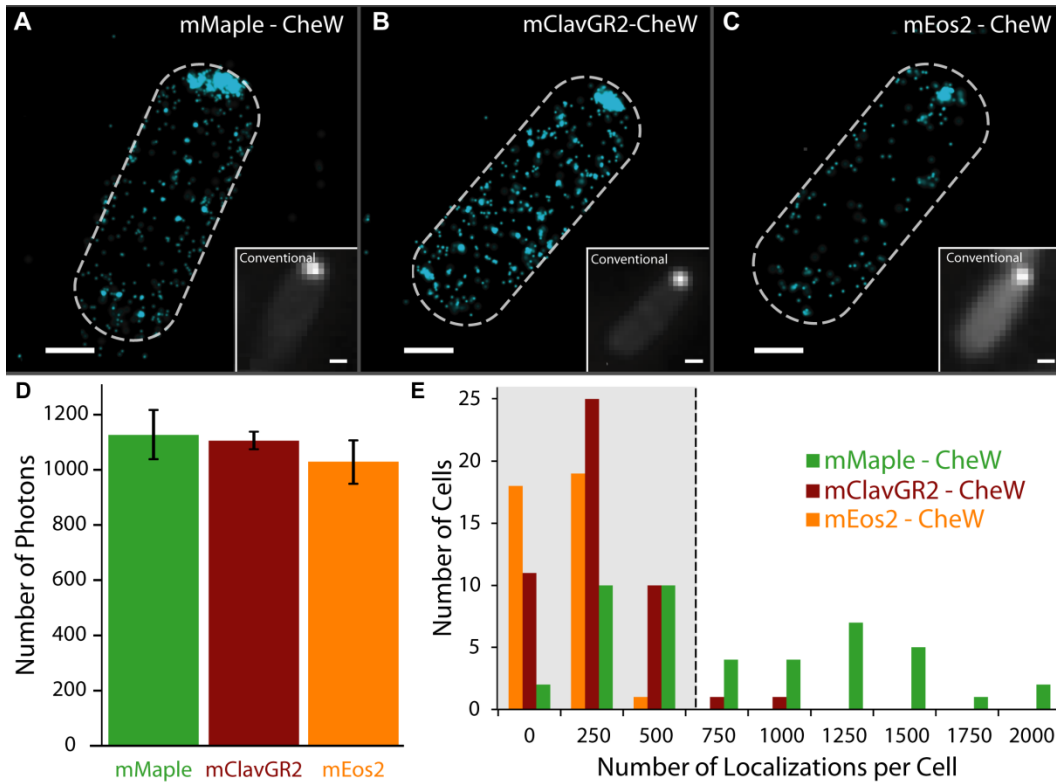


Figure 3.2 (f-)PALM/STORM comparison of mMaple, mClavGR2 and mEos2. (A-C) Images of $\Delta cheW$ *E. coli* expressing CheW fusion proteins at L-arabinose concentrations optimal for function. Images contain (A) 1086 mMaple-CheW localizations (B) 694 mClavGR2-CheW localizations and (C) 229 mEos2-CheW localizations. Scale bars are 500 nm. (D) The mean number of photons emitted by each construct per photoconversion event (error is standard error, $N = 3$ independent measurements from distributions consisting of 4,000 – 32,000 localizations). (E) Distribution of the number of localizations observed for $\Delta cheW$ *E. coli* cells containing CheW fusions to mMaple, mClavGR2, and mEos2. Greater than 96% of cells expressing either mEos2- or mClavGR2-CheW fusions have less than 500 localizations (boxed region), whereas greater than 50% of cells expressing mMaple-CheW fusions have more than 500 localizations. Data collected and image prepared by Ann McEvoy.

In addition to (f-)PALM/STORM, the exceptional brightness of mMaple's green state allows its imaging with SIM (Figure 3.3). In comparison, EGFP-CheW required a 100× higher induction level to produce SIM images of similar quality, while attempts to acquire analogous data sets with mEos2-CheW failed due to rapid photobleaching of the green state.

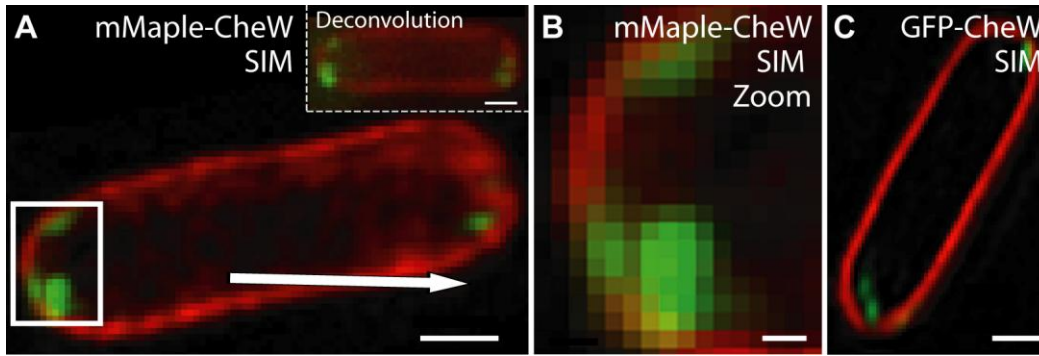


Figure 3.3 mMaple has improved *in vivo* brightness and enables 3D-SIM reconstructions. (A) $\Delta cheW$ *E. coli* expressing mMaple-CheW. (B) A zoom of the polar region of the cell denoted by the boxed region in (A). (C) One 125 nm slice of the 3D-SIM reconstruction of a $\Delta cheW$ *E. coli* expressing GFP-CheW. Red represents fluorescence of the membrane-specific dye FM4-64 and green represents FP fluorescence. Scale bars are 500 nm (A, C) and 100 nm (B). Data collected and image prepared by Ann McEvoy.

The result of the super-resolution imaging experiment suggests that mMaple is a useful “multimodal” FP. However, we were left with the question of what property or properties of mMaple made it so particularly useful in both PALM and SIM imaging. We hypothesized that the exceptional higher numbers of localizations of mMaple compared to other FPs in (f-)PALM/STORM images could be due to improve folding efficiency or more efficient reversible photoswitching of the probe or a combination of both factors. Similarly, the outstanding performance of mMaple in SIM might be attributable to improve folding efficiency or enhanced photostability or both. In this chapter we describe a series of biophysical and biochemical characterizations to seek answers for the above questions. Where appropriate, we have indicated experiments that were performed by our collaborators Ann McEvoy (UC Berkeley), Michael Davidson (Florida State University), and the group of Helge Ewers (ETH Zurich).

3.2 Results and discussion

3.2.1 *In vivo* folding and maturation of pcFPs

Measurements of the rate of chromophore maturation at 37°C showed that mEos2 has much faster maturation kinetics than both mMaple and mClavGR2. Indeed under our experimental conditions, mEos2 is fully matured before the initiation of the measurement (Figure 3.4A). In spite of the slower maturation kinetics, SDS-PAGE reveals that the protein expression levels of mMaple and mClavGR2 in *E. coli* are about 6× higher than mEos2 (Figure 3.4B and C). In addition, both mClavGR2 and mMaple are found 100% in the soluble fraction. In contrast, a significant fraction of mEos2 (29%) existed in the inclusion body, suggesting defective folding. Hence the 10× increase number of localization observed for mMaple in (f-)PALM/STORM images is at least partially due to the increased amount of properly folded protein.

The improvement on the folding efficiency of mMaple should also facilitate the overall folding of the chimeric protein when it is fused to other partners. We performed similar SDS-PAGE analysis for *cheW* knock-out *E. coli* expressing the FP-CheW constructs. The induction levels were previously optimized by Ann McEvoy to yield comparable functionality recovery for CheW. Accordingly, we observed similar protein expression levels for different constructs (Figure 3.5A). Nonetheless, despite the fact that both GFP and mEos2 have better intrinsic brightness than mMaple, the fluorescence intensity of the GFP-CheW and mEos2-CheW is substantially dimmer than that of mMaple-CheW (Figure 3.5B and C), suggesting a substantial amount of non-productive protein misfolding in these variants. The increased fraction of fluorescent mMaple-CheW most likely accounts for the increased number of localizations in (f-)PALM/STORM, as well as the relatively low induction levels needed to provide strong green-state fluorescence signal.

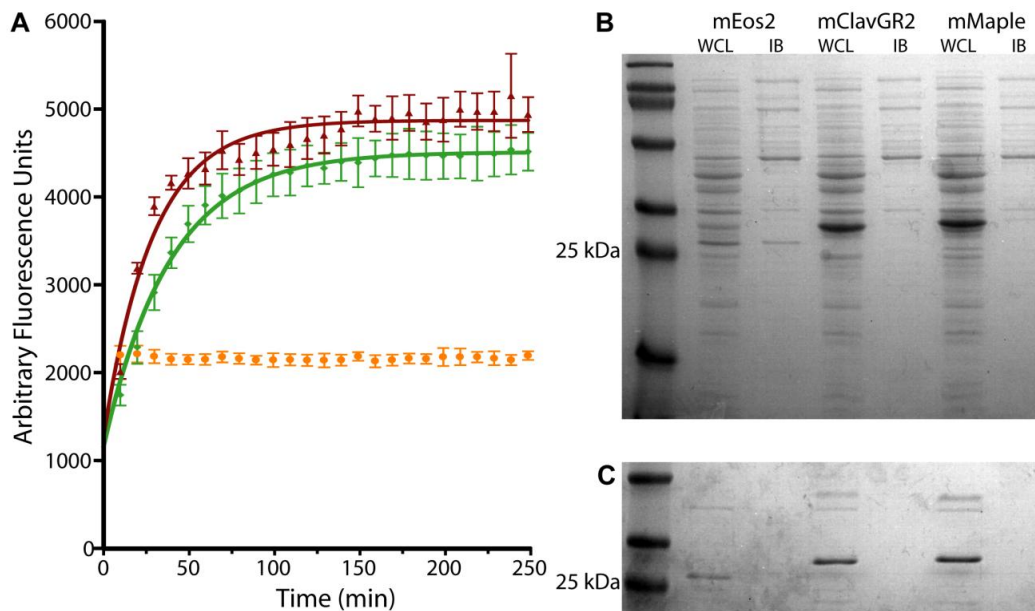


Figure 3.4 Expression and maturation of pcFPs in *E. coli*. (A) Maturation of mClavGR variants and mEos2 at 37 °C. The maturation profiles of mMaple (green) and mClavGR2 (dark red) can be fit as monoexponential curves with time constants of 39 min and 29 min, respectively. Under the conditions of this experiment, mEos2 (orange) is approximately 50% as bright as mMaple and, appears to have fully matured prior to the initial measurement. Each curve represents the average of six independent measurements and error bars represent standard deviations. (B) SDS-PAGE of the soluble and insoluble fractions of *E. coli* expressing pcFPs described in this work. Lane 1 is the protein ladder. For each construct, one lane corresponds to the whole cell lysate (WCL) and the other lane corresponds to protein from inclusion bodies (IB). The relative intensity of FP bands in the WCL and IB fractions, respectively, are: 17 and 7 for mEos2; 100 and 0 for mClavGR2; 113 and 0 for mMaple. Overall, mMaple shows the highest expression and folding efficiency with 100% of the protein in the soluble fraction, while mEos2 has the lowest expression and folding efficiency with 29% of the total expressed protein located in the IB fraction. (C) The same samples as in (B), following purification by Ni²⁺/NTA affinity chromatography. Figure prepared in collaboration with Ann McEvoy.

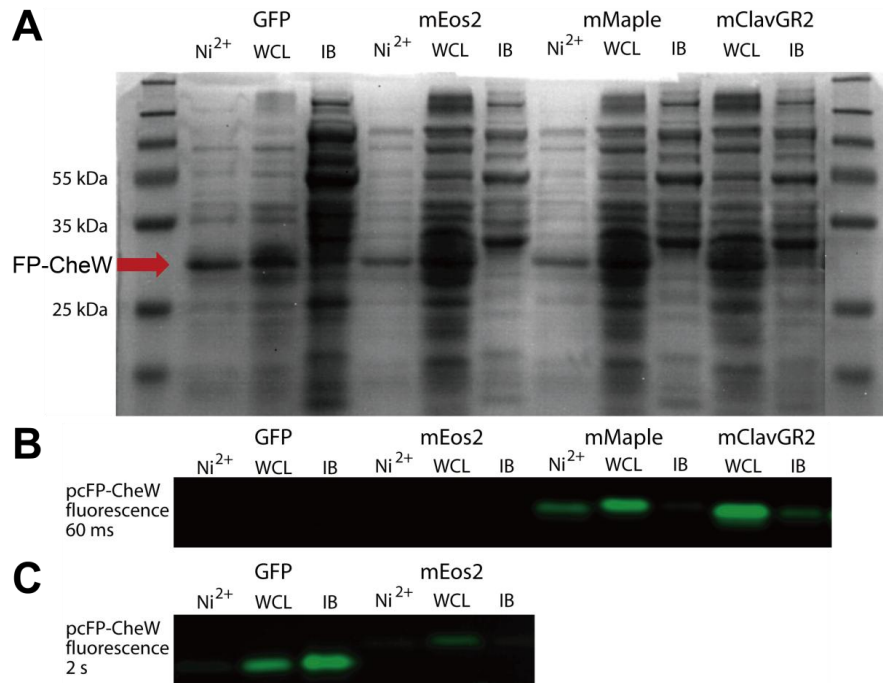


Figure 3.5 (A) Pseudo native SDS-PAGE gel of the soluble and insoluble fractions of *E. coli* expressing pcFP-CheWs (row denoted by red arrow) described in this work. First and last lanes are the protein ladder. Proteins were either purified from the soluble fraction by Ni²⁺-NTA, solubilized from inclusion bodies with urea (IB), or loaded onto the gel as a whole cell lysate (WCL). Relative intensity of the bands is: 132, 120 and 27 for EGFP-CheW; 82, 52 and 0 for mEos2-CheW; 100, 164 and 0 for mMaple-CheW; 154 and 0 for mClavGR2; Samples were not denatured prior to loading. (B-C) Fluorescence image of the bands corresponding to pcFP-CheW fusions extracted from *E. coli* and analyzed by SDS-PAGE gel. Exposure times were increased from 60 ms to 2 sec in (C) in order to visualize the fluorescence of GFP-CheW and mEos2-CheW. Relative fluorescence intensities after correction for exposure times are: 1.00, 2.56, 0 (mMaple-CheW), 4.79, 0.43 (mClavGR2-CheW), 0, 0.08, 0.14 (EGFP-CheW), 0, 0.03 and 0 (mEos2-CheW). Figure prepared in collaboration with Ann McEvoy.

3.2.2 Reversible photoswitching

One possible explanation for the higher numbers of localizations of mMaple observed in (f-)PALM/STORM images would be that mMaple has a higher propensity for reversible photoswitching, which would increase the number of observed localizations per cell. For instance, the red fluorescent state of mEos2 can undergo multiple cycles of reversible

photoswitching into a long-lived dark state [209]. This effect is important to consider for accurate protein counting measurements using (f-)PALM/STORM images [209-211]. Indeed, we found that the red state of mMaple and mClavGR2 also exhibits reversible photoswitching (Figure 3.6). When the photoconverted protein is exposed to intense 532 nm laser light, the protein is switched from the red fluorescent state to the dark state (Figure 3.6A), which can be switched back into the fluorescent state by either 460 nm light, 405 nm light or white light (Figure 3.6B). However, no switching back to the red state was observed if the protein was kept in the dark.

To quantify the extent of reversible photoswitching under (f-)PALM/STORM imaging conditions, Ann used a previously described clustering algorithm [210] to group closely spaced protein localizations (i.e. <30 nm interlocalization spacing), and classified these groups as resulting from reversible switching of a single protein. Analysis of the (f-)PALM/STORM images of *E. coli* with cytoplasmically expressed pcFPs revealed that 56% of mMaple, and over 80% of both mClavGR2 and mEos2 localizations did not have a second localization within 30 nm (Figure B1 in Appendix B). Even so, counting by cluster instead of by localization still found 765 ± 283 of mMaple proteins per cell, 7× more than mEos2 (109 ± 55).

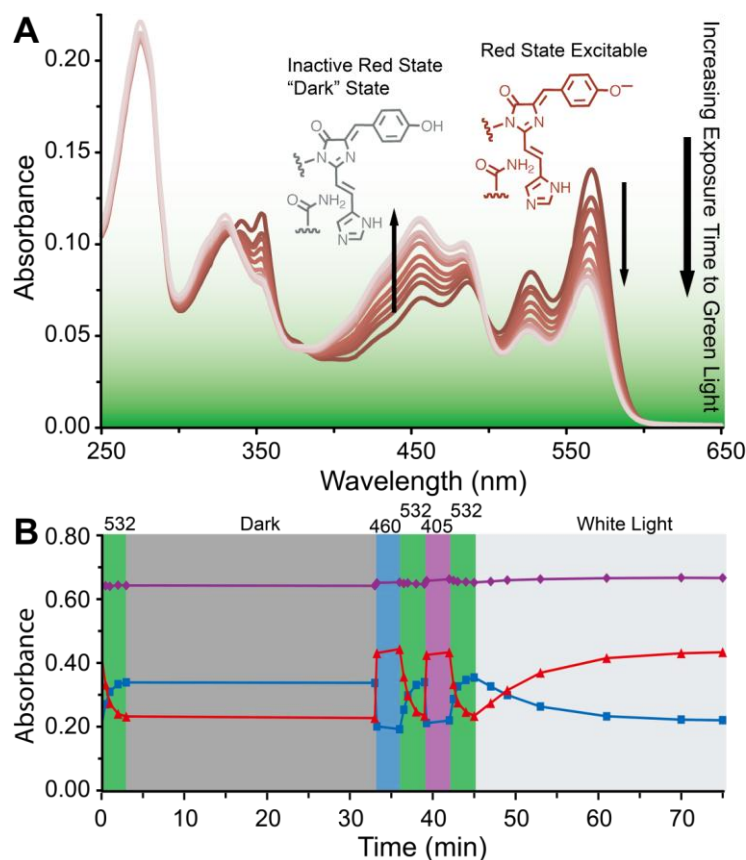


Figure 3.6 Reversible photoswitching of photoconverted (red) mMaple. (A) Photoconverted mMaple can be further photoswitched to a “dark” non-fluorescent state by illumination with green light. (B) The dark state of the photoconverted mMaple reversibly photoswitches back to the red fluorescent state as seen by changes in the absorbance spectra in response to different light sources. The absorbance at 280 nm (purple line, corresponding to total protein concentration), the absorbance at 457 nm (blue line, corresponding to the dark post-conversion red state) and the absorbance at 566 nm (red line, corresponding to the red fluorescent state of the protein) are plotted. Similar results were obtained for mClavGR2. Figure prepared in collaboration with Ann McEvoy.

To further quantify possible reversible switching events, biotinylated single mMaple and mEos2 proteins were immobilized on a glass coverslip and imaged using (f-)PALM/STORM (Figure 3.7). During image analysis, closely spaced localizations were once again grouped into clusters. Cluster size distributions for mMaple, mEos2, and the negative control for which no fluorescent proteins were present is shown in Figure 3.7C. The background subtracted cluster size distributions for each pcFP revealed

that approximately 35% of mMaple localizations and 65% of mEos2 localizations were observed as single localizations (Figure 3.7D). Therefore we conclude that, under our imaging conditions, the red fluorescent state of mMaple has a two-fold higher propensity to reversibly photoswitch than mEos2. This two-fold increase tendency of reversible photoswitching does not fully account for the 7× increased number of localization found in *E. coli* expressing mMaple.

We next investigated whether differences in photoconversion probability could explain the increased number of localizations observed for mMaple in (f-)PALM/STORM. The two factors that contribute to the photoconversion probability are the extinction coefficient at 405 nm and the quantum yield for photoconversion. The extinction coefficient of mMaple at 405 nm is $10,300 \text{ M}^{-1}\text{cm}^{-1}$, whereas mEos2 is $4,400 \text{ M}^{-1}\text{cm}^{-1}$. However, this 2.3× higher absorbance for mMaple is counterbalanced by a 5.5× lower quantum yield of photoconversion (see Section 3.4.2 for calculation). Taking both factors into account, mEos2 has a 2.4× higher probability of photoconverting under identical illumination conditions. Accordingly, the increased number of localizations obtained with mMaple is not due to differences in photoconversion probability.

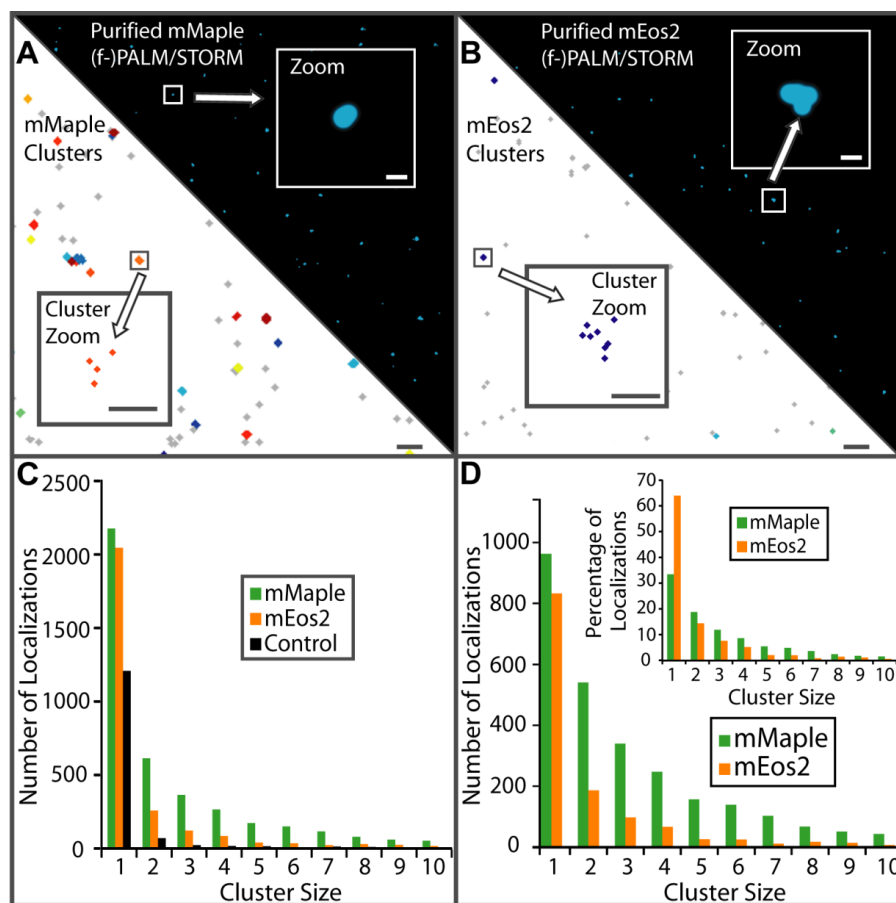


Figure 3.7 (f-)PALM/STORM analysis of purified pcFPs. (A-B) Composite images of purified (A) mMaple and (B) mEos2 proteins immobilized on a coverslip. Localizations are represented as normalized 2D Gaussian peaks (right half) and represented as single localizations and clustered markers (<30 nm interlocalization spacing) (left). Scale bars are 2 μ m and 50 nm (zooms). (C) Cluster size distribution for purified mMaple and mEos2 as well as a no protein control. (D) False-positive corrected cluster size distributions for purified mMaple and mEos2 demonstrate that mMaple is approximately twice as likely to reactivate as mEos2. (inset) Percentage of false-positive corrected localizations found to be in clusters for purified pcFPs. 35% of mMaple and 65% of mEos2 do not have a second localization within 30 nm. Data collected and image prepared by Ann McEvoy.

3.2.3 Photostability study of pcFPs

Obtaining SIM reconstructions, particularly 3D-SIM reconstructions, typically requires a large number of images of a sample, therefore it is important to use bright and photostable fluorophores to increase image

quality. The increased folding efficiency of mMaple certainly provides partial explanation for its superior performance in SIM imaging. However, the apparent lack of appreciable fading during prolonged SIM image acquisition also suggested improved photostability for mMaple's green state. In collaboration with Paula Cranfill and Michael Davidson at the National High Magnetic Field Laboratory (US), we quantified the pcFPs' resistance to photobleaching in live cells. Under widefield illumination, mMaple's green state was 14-fold more photostable than mEos2's green state, as judged by time to 50% loss of signal (65.1 s for mMaple and 4.6 s for mEos2) (Figure 3.8 and Table 3.1). Therefore, the high quality of the *E. coli* SIM images is attributable to mMaple's higher *in vivo* green state photostability. Photobleaching of the red state of mMaple under widefield illumination is slightly faster than mClavGR2 and mEos2 (Figure 3.8) but 2.4x slower than mEos2 under confocal illumination (Figure 3.9). The photostability advantage of the green state of mMaple under confocal illumination is less obvious, because all of the three pcFPs are rapidly photobleached (Figure 3.9). Values of the photostability are tabulated in Table 3.1. We have previously reported in Chapter 2 the widefield (at 202.6 mW/cm² and 489.9 mW/cm² for green and red, respectively) and confocal (at 100 μW for both green and red) photobleaching times for mClavGR2 and mEos2. It should be clarified that the relative order of photobleaching half-times for different variants can change as a function of power. Accordingly, the numbers provided here are only relevant for the output power at which they were measured.

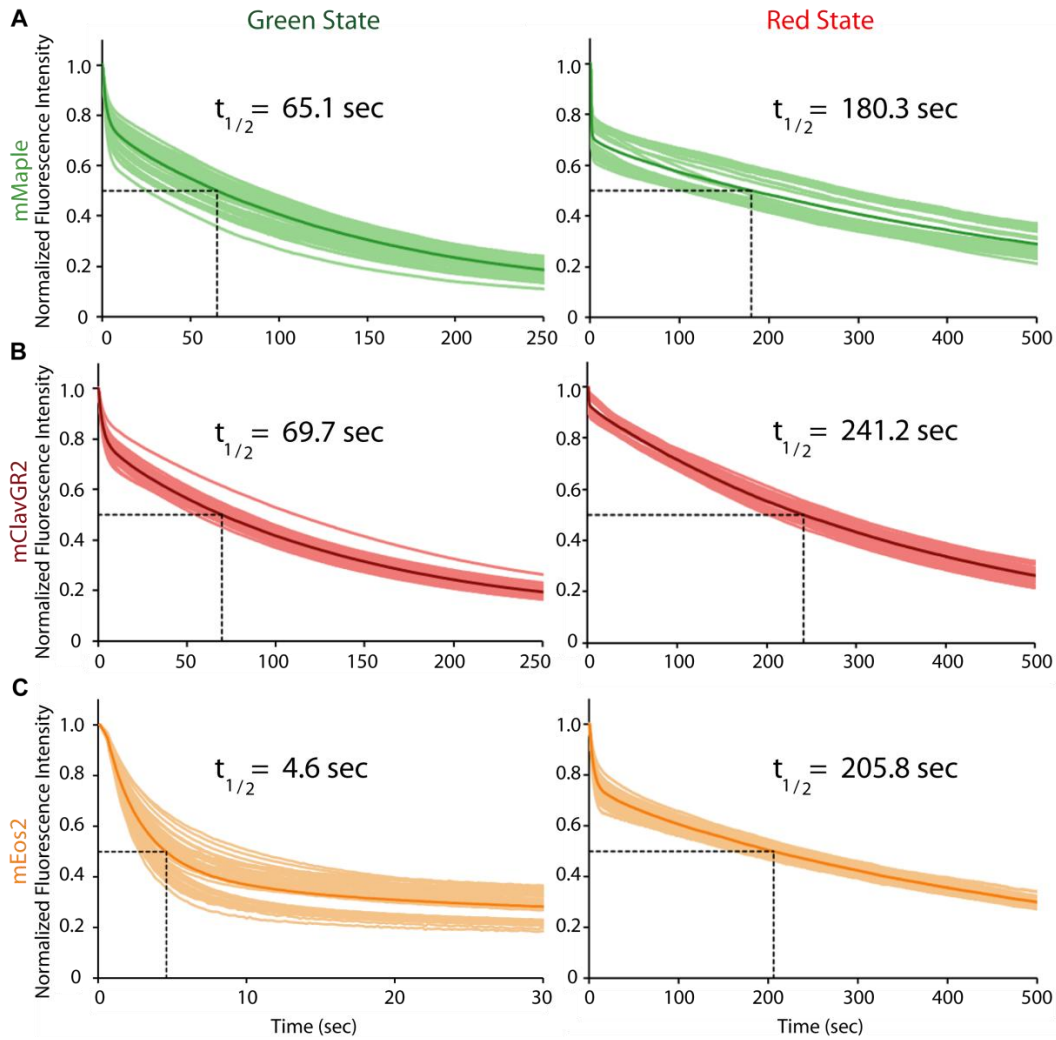


Figure 3.8 Widefield imaging photobleaching curves for pcFPs. Photobleaching curves of the green state (left panels) and the red state (right panels) of pcFP-H2B fusions expressed in HeLa S3 cells with widefield illumination. Each curve represents the photobleaching behavior of an individual cell and the darker colored curve is the average. Average time when the fluorescence intensity of the green states decreased to half of the initial intensity are 65.1 s for mMaple (A) (45 cells), 69.7 s for mClavGR2 (B) (42 cells), and 4.6 s for mEos2 (C) (48 cells). The average time for the red state is 180.3 s for mMaple (A) (36 cells), 241.2 s for mClavGR2 (B) (35 cells), and 205.8 s for mEos2 (C) (49 cells). Values are tabulated in Table 3.1. Photobleaching experiments were performed by Paula Cranfill and Michael Davidson. Figure prepared in collaboration with Ann McEvoy.

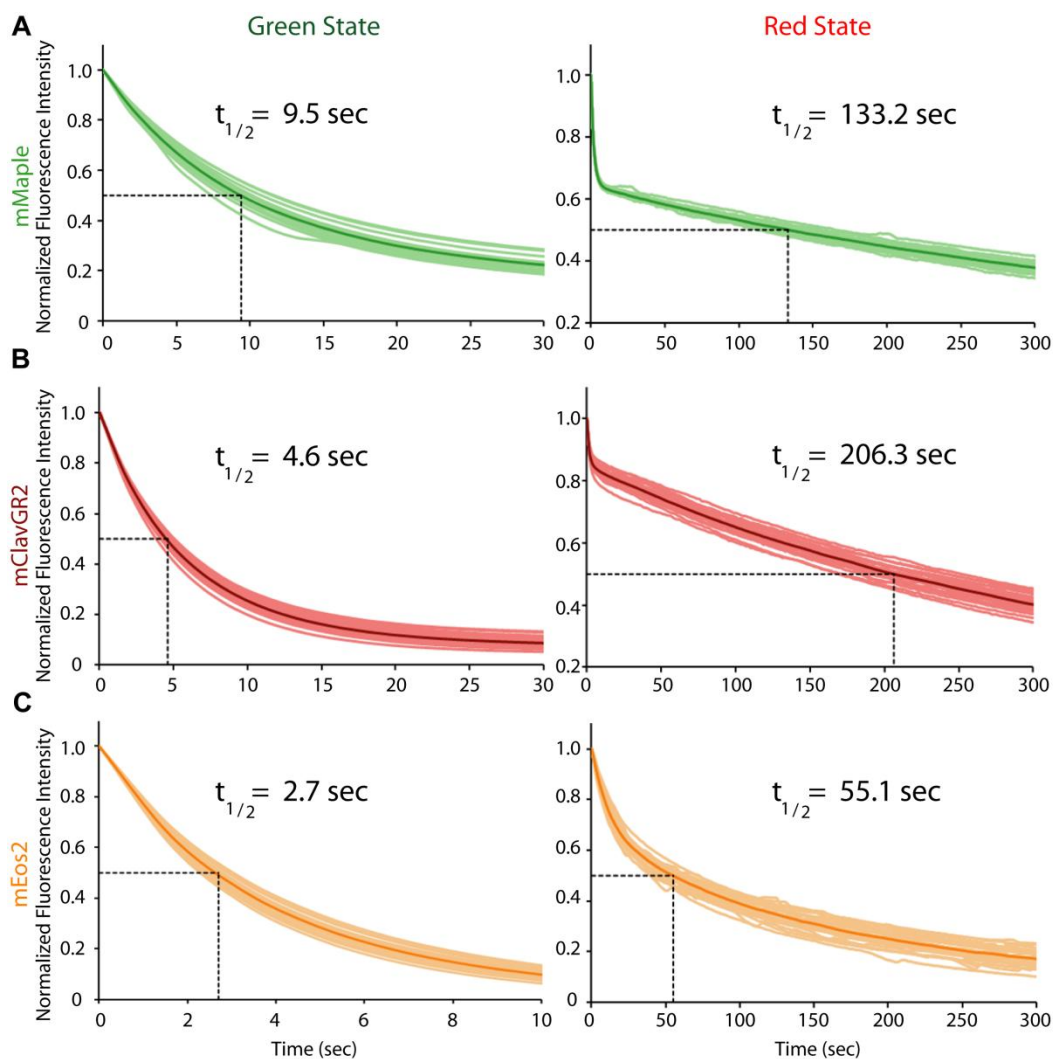


Figure 3.9 Confocal imaging photobleaching curves for pcFPs. Photobleaching curves of the green states (left panels) and the red states (right panels) of pcFP-H2B fusions expressed in HeLa S3 cells with confocal illumination. Each curve represents the photobleaching behavior of an individual cell and the dark colored curve is the average. Average time for the green state fluorescence intensity to decrease to half of the initial intensity are 9.4 s for mMaple (A) (22 cells), 4.6 s for mClavGR2 (B) (30 cells), and 2.7 s for mEos2 (C) (28 cells). The average time for the red state fluorescence intensity to decrease by half is 133.2 s for mMaple (A) (27 cells), 206.3 s for mClavGR2 (B) (23 cells) and 55.1 s for mEos2 (C) (25 cells). The rapid initial bleaching in mMaple and mClavGR2 is likely attributable to bright-to-dark photoswitching. Values are tabulated in Table 3.1. Photobleaching experiments were performed by Paula Cranfill and Michael Davidson. Figure prepared in collaboration with Ann McEvoy.

Table 3.1 Photobleaching rates for pcFP-H2B fusions in live cells

Protein	State	Time to bleach to 50% of initial intensity (s)	
		Widefield ^a	Confocal ^b
mMaple	green	65.1	9.4
	red	180.3	133.2
mClavGR2	green	69.7	4.6
	red	241.2	206.3
mEos2	green	4.6	2.7
	red	205.8	55.1

^a Widefield photobleaching of both states was performed at an output power of 11.4 mW/cm². ^b Confocal photobleaching of both states was performed at an output power of 120 μ W. All photobleaching curves are provided in Figure 3.8 and Figure 3.9.

3.2.4 mMaple's performance in mammalian cells

We also tried to determine whether the folding advantage of mMaple displayed in *E. coli* can be carried on to mammalian cells. We fused the pcFPs to cytoskeleton proteins such as actin and actinin and expressed them in HeLa cells. Widefield imaging reveals that all constructs display correct localization patterns of actin (Figure 3.10A-C) or actinin (Figure 3.10F-H) and a similar overall brightness in the green state was observed. We next use flow cytometer to statistically quantify the fluorescent intensity of each cell. For actin fusions, the overall distribution of mEos2 and mClavGR2 is slightly shifted to a higher brightness and the relative median fluorescence intensity after excitation and emission correction is: 0.68 for mMaple, 0.77 for mClavGR2 and 1.00 for mEos2 (Figure 3.10D). For the actinin fusions, overall distribution of mEos2 is slightly dimmer than mClavGR2 and mMaple and the relative corrected median fluorescence intensity is: 1.03 for mMaple, 1.25 for mClavGR2 and 1.00 for mEos2 (Figure 3.10E). Neither mMaple-actin nor mMaple-actinin is substantially

brighter than the corresponding mEos2 or mClavGR2 fusions. Indeed, none of the pcFPs display absolute folding or fusion advantage when expressed in mammalian cells. This is not unexpected since protein expression levels in mammalian cells are affected by many factors such as the cell line used, transfection efficiency and cell health condition and others.

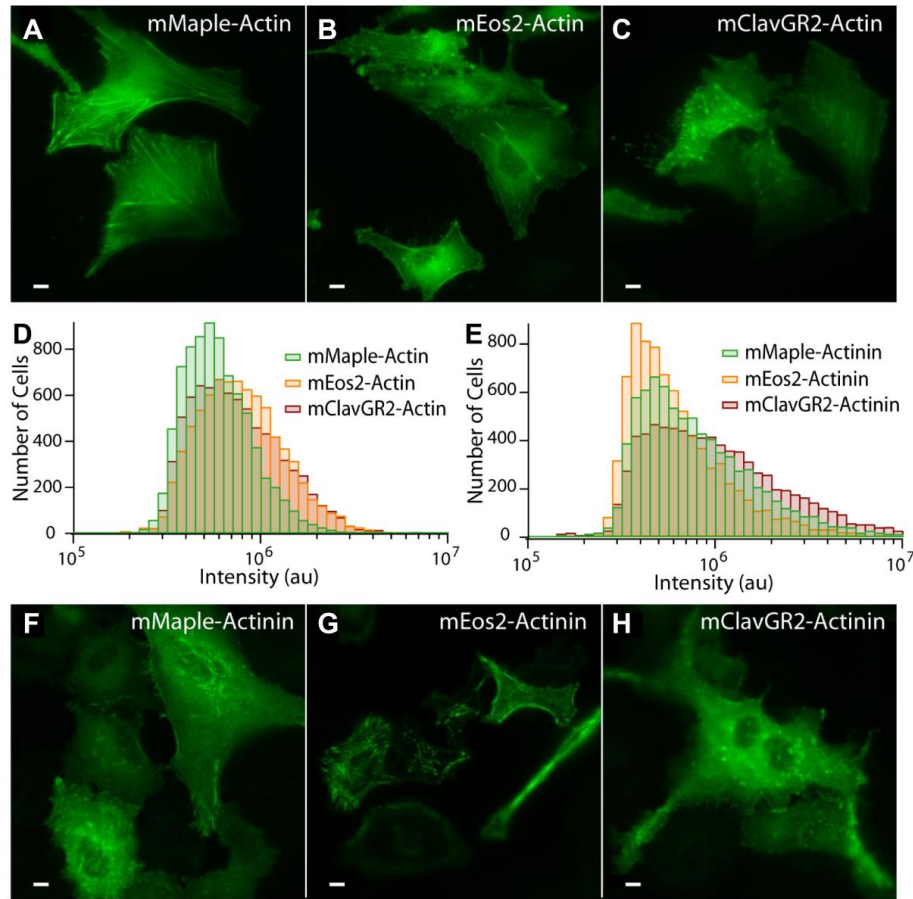


Figure 3.10 pcFP fusions expressed in mammalian cells. (A-C) Representative widefield fluorescence images of HeLa cells transfected with plasmids encoding either mMaple-actin (A), mEos2-actin (B), or mClavGR2-actin (C). (D) Flow cytometric analysis of the green fluorescence for HeLa cells transfected with the same plasmids used in A-C. A total of 7000 fluorescent cells were analyzed in each experiment. (F-H) Representative images of HeLa cells transfected with actinin fusions of mMaple (F), mEos2 (G), and mClavGR2 (H). (E) Flow cytometric analysis of the green fluorescence from HeLa cells transfected with the plasmids used in F-H. Figure prepared in collaboration with Ann McEvoy.

Further examination of the actin constructs showed that though the overall brightness of the mMaple-actin constructs was not increased relative to the other pcFPs, the photostability of mMaple-actin was improved by almost three-fold (10.0 s for mMaple in comparison to 3.6 s as judged by confocal imaging) (Figure B2 in Appendix B), increasing the utility of the mMaple-actin construct in applications requiring green-state photostability.

3.3 Conclusions

Our collaborator Ann McEvoy has demonstrated that mMaple can be used as a multi-modal probe for both (f-)PALM/STORM and SIM and this is the first example for a same FP to be successfully imaged by two different super-resolution microscopy techniques. Through a series of biochemical and biophysical characterization, we determined that this beneficial property results from a combination of a significantly improved protein folding efficiency and improved photostability of the green state. The advantages and limitations of mMaple relative to mEos2, for specific applications and uses described in this work, are summarized in Table 3.2. We expect that this new variant will be of utility as it allows the dynamics of protein complexes to be characterized in live cells with increased spatial resolution and good temporal resolution (e.g. by SIM or conventional imaging), and also allows the precise localization of the same fusion proteins to be determined with (f-)PALM/STORM. Also, we hope the detailed characterization described here could serve as a reference and a reminder for potential pitfalls for researches who are interested in using mMaple or other pcFPs to decipher biological phenomena.

Table 3.2 Advantages and limitations of mMaple

Imaging Method	mMaple Properties
Widefield epi-fluorescence, green state	~14x higher photostability than mEos2
Widefield epi-fluorescence, red state	Similar photostability to mEos2
Confocal, green state	~3.5x higher photostability to mEos2
Confocal, red state	~2.4x higher photostability to mEos2
SIM, green state	Higher green-state photostability and higher fraction of properly folded proteins enhance image quality for <i>E. coli</i>
Single-molecule localization, (f-)PALM/STORM	Similar localization precision to mEos2. Twice as likely to reversibly photoswitch as mEos2.

3.4 Materials and methods

3.4.1 General methods and materials

See Section 2.4.1. Mass spectrometry (MS) was performed by the MS Facility in the Department of Chemistry, University of Alberta.

3.4.2 Protein purification and characterization

Proteins were purified as described in Section 2.4.4.

The quantum yield of photoconversion (Φ_{PC}) was determined using $\Phi_{PC} = [\text{red}]_{1\text{min}} / \text{Int}[I(x) \times (1 - 10^{-\text{abs}(x)}) dx]$, where $[\text{red}]_{1\text{min}}$ is the concentration of red species after 1 min of photoconversion with the 405 nm LED, $\text{abs}(x)$ and $I(x)$ are the sample absorbance and illuminating light intensity at wavelength x , respectively, and the integral runs from 390 to 420 nm. The emission profile of the 405 nm LED used for photoconversion is defined as $I(x) = I_0 \exp(-(x-405)^2/36)$ based on the emission maximum and the full width at half maximum (FWHM) provided in the product catalogue.

To measure the maturation profiles of mMaple, mClavGR2 and mEos2, *E. coli* transformed with pBAD/His B plasmids bearing the encoding cDNA

was cultured overnight. The culture was diluted to an OD₆₀₀ of 0.6, purged with argon for 20 min, sealed with a rubber septum, and incubated for another 1 h to allow thorough consumption of the residue oxygen. L-arabinose (0.025%) was then added via a syringe with needle to induce expression of the FPs. After 4 h of incubation with shaking at 37 °C, the cultures were transferred to an ice bath for 10 min and maintained at 4 °C. Cells were centrifuged at 8000 rpm for 2 min, and cell pellets were lysed using vacuum-degassed B-PER II (Pierce) and incubated at room temperature for 10 min. The lysate was then centrifuged at 12,500 rpm for 5 min, and the supernatant was diluted 10× into PBS (pH 7.4). Fluorescence maturation was monitored at 37 °C using a Safire2 plate reader (Tecan).

To evaluate the expression level of mMaple, mClavGR2 and mEos2, *E. coli* transformed with pBAD/His B plasmids bearing the encoding cDNA was cultured overnight. The culture was diluted to an OD₆₀₀ of 0.6 in a total volume of 4 mL and L-arabinose (final concentration 0.02%) was added to induce the expression of the FPs. After 2 h growing at 37 °C with shaking, the cells were collected, lysed with 50 µL of B-PER (Pierce), and centrifuged to separate the soluble proteins from insoluble material. An aliquot of the supernatant was taken for later SDS-PAGE analysis. The cell pellet was rinsed once with B-PER and once with 2 M urea. The pellet was then re-dissolved in 100 µL 8 M urea and centrifuged at 12000 rpm for 10 min. An aliquot of the supernatant was taken for later SDS-PAGE analysis. Both the supernatant from the lysate and the supernatant from the pellet extract were further purified by Ni²⁺-NTA-conjugated beads and all samples were analyzed by SDS-PAGE. The intensity of the bands was analyzed using ImageQuant RT ECL (GE Health).

Evaluation of the expression level of FP-CheW fusions was performed in a similar way excepted that (1) A *cheW* knock-out *E. coli* strain was used; (2) cells were grown overnight in T-broth (1% Tryptone, 0.5% NaCl, pH 7.0) supplemented with ampicillin at 30°C with aeration. Day cultures were

inoculated into T-broth at an OD₆₀₀ of approximately 0.01 and grown at 30°C until they reached an OD₆₀₀ of 0.1 – 0.3. Protein expression was induced by adding 0.01% or 0.02% L-arabinose for 3 hrs, as indicated. These growth conditions is exactly same to those used to grow *E. coli* cells for superresolution imaging; (3) 40 µg of each whole cell lysate was loaded onto the gel. Protein concentrations were determined using the BCA kit (Thermo Scientific). Fluorescent image of the gel were acquired using a custom built imaging system equipped with a 300 W Xeon lamp, a 450-490 nm filter for excitation and a 500-550 nm filter for emission. Fluorescent intensity was analyzed using Image Pro Plus (Media Cybernetics).

Purified mMaple and mEos2 were as described above and then biotinylated using EZ-Link Sulfo-NHS-Biotin (Pierce) following the manufacturer's instruction. This biotinylation kit would provide a spacer about 22.4 Å between the pcFP and the biotin module. The spectral profile of absorbance, excitation and emission of the pcFPs were measured and found to remain unchanged. The number of labeled biotins per pcFP molecule was determined to be a distribution that ranged from 7-13, based on MALDI-MS.

3.4.3 Live cell imaging

Construction of mammalian expression plasmid. To create the pcFP-actin constructs, we amplified the gene encoding pcFP with a 5' primer with an *NheI* site and a 3' primer with an *XhoI* site. The purified PCR products were then digested and ligated into pEGFP-actin (Clontech), whose FP-coding gene has been previously removed by the same restriction enzymes. pcFP-actinin constructs were generated in a similar way using a 5' primer encoding an *AgeI* site, a 3' primer encoding a *BspEI*, and the pactinin-EGFP vector (Clontech). Plasmid DNA for transfection was prepared using the GeneJET™ Plasmid Miniprep Kit (Fermentas)

Mammalian cells maintenance and transient transfection. HeLa cells (CCL2 line; ATCC) were maintained in Dulbecco's Modified Eagle Medium (DMEM) (Invitrogen) supplemented with 10% fetal bovine serum (FBS) (Sigma), 2 mM GlutaMax (Invitrogen) and penicillin-streptomycin. Transfection was carried using TurboFect™ (Fermentas) according to the manufacturer's protocol. Imaging or cytometry experiment was conducted 36~48 h after transfection. The growing medium was changed into HEPES (25mM) buffered Hanks' Balanced Salt Solution (HHBSS) for imaging. Imaging was carried on an inverted Nikon Eclipse Ti microscope equipped with a 150W Lumen 200 metal halide lamp (Prior Scientific), a 60× oil immersion objective (Nikon), and a 16-bit 512SC QuantEM EMCCD (Photometrics).

For cytometry experiment, cells were trypsinized, collected by centrifugation, resuspended in HHBSS and analyzed on a C6 Flow Cytometer (Accuri). About 7000 to 20000 events were recorded, depending on the transfection efficiency, but only the first 7000 events were plot in Figure 3.10. At least two replicas were conducted for each constructs. Brightness corrected for excitation and emission was calculated via $F_{\text{corr}} = F_{\text{app}} / \text{relative brightness} = F_{\text{app}} / (EC_{488} \times \alpha)$, where F_{app} is the average of the acquired median fluorescence intensity, EC_{488} is the FP's extinction coefficient at 488 nm, and α is the fraction of the emission spectrum passing the given filter (515-545 nm).

Chapter 4 Development of highlightable Ca²⁺ indicators for live cell imaging⁴

⁴ A version of this chapter has been accepted for publication on *Journal of the American Chemical Society* (Hoi, H., et al. *Journal of the American Chemical Society* 2012 DOI: 10.1021/ja310184a).

4.1 Introduction

The calcium ion (Ca^{2+}) is a ubiquitous intracellular signaling molecule involved in numerous cellular processes. For example, Ca^{2+} plays regulation roles during cell fertilization, differentiation, heart muscle cell contraction, and the processes of learning and memory, to name just a select few [212]. The cytoplasmic concentration of Ca^{2+} in a resting (unstimulated) cell is around 100 nM and can increase to 1000 nM or more upon stimulated release from intracellular stores. This relatively simple process of elevating the cytoplasmic concentration, when combined with modulation of oscillation frequency and/or amplitude and spatially confined subcellular location, create distinct spatial and temporal patterns that can trigger a variety of cellular responses. The robust and versatile character of Ca^{2+} signaling is made possible by an extensive molecular repertoire of signaling components consisting of ion channels, cell surface receptors, kinases, and various Ca^{2+} binding proteins. There is a tremendous amount of interest, extending from basic science to applied health research, in gaining a better understanding of how this versatile Ca^{2+} signaling system functions in specific cellular processes. One of the single most important tools in this area of research is the technique of optical Ca^{2+} imaging [213]. By enabling researchers to optically observe the real time dynamics of Ca^{2+} inside cells, tissues, or whole organisms, Ca^{2+} imaging accelerates the rate at which we can come to understand the mechanisms and consequences of processes that rely on Ca^{2+} -dependent signaling.

An essential class of tools for single cell Ca^{2+} imaging is the optical Ca^{2+} indicators. The current arsenal of Ca^{2+} indicators includes designs based on synthetic organic dyes, bioluminescent proteins such as Aequorin, and engineered fluorescent proteins (FPs) [214]. As FP-based Ca^{2+} indicators are genetically encodable, they can be introduced into cells using the minimally invasive method of plasmid DNA transfection. Relative to synthetic dyes, FP-based indicators have the additional advantages of specified targeting to subcellular compartments and expression in

transgenic organisms [215]. Unlike Aequorin, FP-based Ca^{2+} indicators do not require externally provided substrates to yield luminescence. While the current toolbox of Ca^{2+} indicators provides a variety of choices with respect to indicator class (genetic vs. synthetic), fluorescence hue, and Ca^{2+} affinity, there is a continued demand for new tools with improved or unprecedented properties [214, 216]. One desirable property of a Ca^{2+} indicator would be the ability to highlight specific experimenter-chosen cells within a tissue such that their Ca^{2+} dynamics could be imaged in a distinct spectral channel. This would enable researchers to spectrally isolate and image the activity of specific cells in a transgenic animal with ubiquitous expression of a suitable genetically encoded indicator. Similarly, irreversible highlighting would allow researchers to return to, and image the activity of, a cell imaged in an earlier experiment.

In an effort to create a highlightable Ca^{2+} indicator for live cell imaging we have attempted to use protein engineering to combine the properties of a well-established class of FP-based Ca^{2+} indicator with the properties of four previously reported classes of photoactivatable (pa) or photoconvertible (pc) FPs. The class of FP-based Ca^{2+} indicator used in this work is the single FP design common to the Pericam and G-CaMP series of indicators [164, 165, 217, 218]. In this design, the calcium-binding protein calmodulin (CaM) and its binding peptide M13 (from skeletal muscle myosin light-chain kinase) are genetically fused to the N- and C- termini of a circularly permuted FP (cpFP). Conformational changes resulting from binding of Ca^{2+} by CaM lead to a change of the FP chromophore environment that increases its fluorescence [166, 167]. Our lab has recently reported a method for high throughput screening of large libraries of G-CaMP variants for identification of mutants with improved function or altered fluorescence hue [168]. This effort produced the GE CO (abbreviation of genetically encoded Ca $^{2+}$ optical indicator) series of multicolored Ca^{2+} indicators that serve as the templates for this work.

The four classes of paFP and pcFP explored in this work are: 1) the paGFP-type that increases green fluorescence when illuminated with violet (~400 nm) light [175]; 2) the pamCherry type that increases red fluorescence with violet light [219]; 3) the psmOrange type that converts from orange-to-far-red fluorescent with blue light [220]; and 4) the Kaede-type that converts from green-to-red fluorescent with violet light [75]. We reasoned that introduction of the key mutations responsible for these pa or pc phenotypes into the FP module of appropriate GECOs could produce pa- or pcGECOs.

In this chapter I will describe my work in developing highlighter Ca^{2+} indicators, including their engineering, optimization, characterization and imaging application in live cells.

4.2 Results and discussion

4.2.1 Attempted development of a paGECO based on paGFP and green emissive GECOs

The *Aequorea victoria* GFP (avGFP) undergoes photoactivation (an increase in green fluorescence when excited at 475 nm) when irradiated with UV or ~400 nm light [221]. The chromophore of avGFP exists in equilibrium between the neutral phenol state ($\text{abs}_{\text{max}} = 397 \text{ nm}$) and the anionic phenolate state ($\text{abs}_{\text{max}} = 475 \text{ nm}$). Absorbance by either state results in green fluorescence. However, illumination with UV or ~400 nm light can also lead to decarboxylation of Glu222 (albeit with a lower quantum efficiency) and a rearrangement of the hydrogen-bond network that shifts the chromophore equilibrium towards the phenolate state [222, 223]. With the goal of improving the contrast of this photoactivation and creating a useful FP-based highlighter, Patterson and colleagues screened a library of mutations at position 203. This effort led to the identification of the T203H variant of GFP that has a more than 100× increase in green emission with 488 nm when photoactivated with ~400

nm light [175]. We introduced the analogous V116H mutation into green-emissive GECOs, namely G-GECO1.1-1.3, GEM-GECO1 and GEX-GECO1 [168]. G-GECO1.1-1.3-V116H variants retained strong green fluorescence with 480 nm excitation and no photoactivation was observed after illuminated with 400 nm in *E. coli* colonies. For GEX-GECO1, V116H substantially diminished its brightness, and abolished its Ca²⁺ response. GEM-GECO-V116H had an inverted Ca²⁺ response at ~515 nm when excited at ~ 400 nm. Neither GEX- nor GEM-GECO1 underwent photoactivation even with prolonged photoactivation (26 min). Position 116 (i.e. position 203 in GFP numbering) is located at the interface between the cpGFP and calmodulin in G-CaMP and a threonine at this position was found to participate in a hydrogen bond network which stabilizes the ionized state of the chromophore [166]. Substitution of threonine with valine was introduced during the evolution from G-CaMP2 to G-CaMP3 [218]. We suspect that, in contrast to PA-GFP where the Thr203His mutation stabilizes the neutral chromophore [82], this substitution is stabilizing the anionic form of the chromophore in the context of the G-GECOs.

The E222Q mutation has been reported to make avGFP and its derived variants reversibly photoswitchable [224]. Introduction of this mutation into GEX- and GEM-GECO1 did not introduce an analogous property. No recovery of green fluorescence was observed following photobleaching by illumination with 480 nm light.

4.2.2 Attempted development of a paGECO based on pamCherry and R-GECO

R-GECO1 is a red-emissive GECO based on circularly permuted mApple [168]. Since mApple and mCherry are both derived from *Discosoma* RFP [103], and mCherry has been converted into the photoactivable variant pamCherry [225], we hypothesized that introducing mutations responsible for the photoactivation behavior of pamCherry into

R-GECO1 might generate a photoactivable red fluorescent GECO. It has been proposed the pamCherry initially forms a non-coplanar chromophore structure that lacks the C_α-C_β double bond at Tyr67 that is present in the normal mCherry chromophore [219]. Irradiation promotes the oxidation of the C_α-C_β of the Tyr67 of the chromophore through a mechanism that involves the decarboxylation of Glu215. Mutagenesis studies indicated that Lys70Asn and Ile197Arg are minimally required to the photoactivation. Accordingly, we targeted the equivalent residues in R-GECO, and several others thought to be important for photoactivation (i.e. V148, I165, K167, and L169) for generation of semi-saturation libraries. Libraries were screened in *E. coli* colonies for both photoactivation and Ca²⁺ responses. Several mutants that exhibited photoactivation behavior upon 400 nm light irradiation were identified. Of them, R-GECO-V148M/I165V/K167A, designated as paR-GECO0.1, exhibited the best, though still modest, photocontrast (6-fold). Notably, mutations at Lys70 and Ile197 were not required to make R-GECO photoactivable. paR-GECO0.1 has a single absorbance peak at 402 nm in the visible spectrum range (Figure 4.1A). Irradiation with ~400 nm light generates a new absorbance peak at 565 nm, which emits maximally at 605 nm. Unfortunately, paR-GECO0.1 and other less photoactivable mutants showed only subtle Ca²⁺ responses (15~30% $\Delta F/F_{\min}$) in stark contrast to the 16-fold change of R-GECO1 (Figure 4.1C). Assuming that the photoactivation mechanism of paR-GECO0.1 is similar to that of pamCherry, the post-photoactivated chromophore would be in a *trans* conformation and therefore oriented differently than the *cis* chromophore of mApple and mCherry [219]. Accordingly, the phenolate group of the paR-GECO0.1 chromophore will not be positioned to engage in the key interactions with CaM that give rise to the large fluorescence change of R-GECO1. Discouraged by the modest photocontrast and poor Ca²⁺ response, as well as concerns in the photoactivation mechanism, we did not pursue further directed evolution on paR-GECOs. It is possible that additional directed evolution and

engineering could potentially lead to the development of useful imaging tools based on these variants.

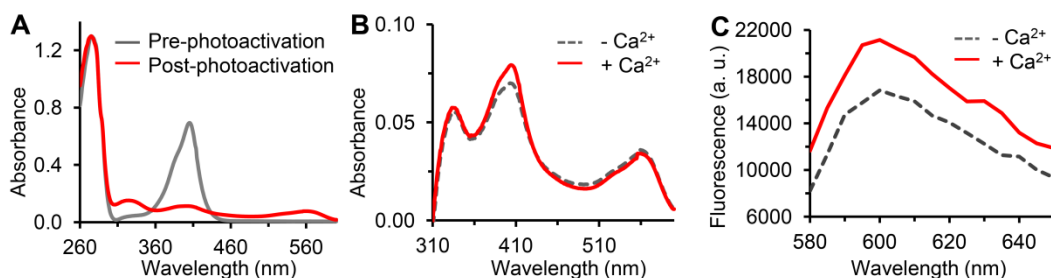


Figure 4.1 Photoactivation and Ca²⁺ response of paR-GECOs. (A) Absorbance spectra of pre- and post-photoactivation states of paR-GECO0.1. Before photoactivation, paR-GECO0.1 has a major absorbance peak at ~400 nm that is weakly blue fluorescent. This peak is photoconverted into a red species after 16 min illumination with 405 nm LED light. (B) Absorbance spectra of the post-photoactivation red species with and without Ca²⁺. (C) Fluorescence spectra of the red species with and without Ca²⁺. Saturating the Ca²⁺-free paR-GECO0.1 with Ca²⁺ causes a 19% increase at the emission maximum.

4.2.3 Attempted development of a pcGECO based on psOrange and O-GECO

PsmOrange is one of the most recently developed photoconvertible FPs [220]. This protein undergoes photoconversion from orange fluorescent into far-red fluorescent with irradiation at blue-green wavelength (458 nm or 488 nm) [220]. Again, we hypothesized that transplanting the key mutations responsible for the photoconversion property into an orange GECO (O-GECO, unpublished work by Jiahui Wu) might generate a pcO-GECO. This O-GECO is a blue-shifted variant of R-GECO with 9 mutations and displaying a remarkable 146x signal increase upon Ca²⁺ binding. Taking suggestions from Benjamin Moeyaert (from the lab of Johan Hofkens at University of Leuven, Department of Chemistry), a visiting graduate student in our lab, we started with targeting position 221 in O-GECO (equivalent to position 63 in psmOrange, position 64 in mMaple). This position is potentially important since all photoconvertible FPs have a hydrophobic amino acid residue here, whereas O-GECO has

glutamine at this position. The resulted O-GECO-Q221L was extremely dim, only displaying fluorescence after two weeks incubation at room temperature. Subsequent illumination with the O-GECO-Q221L expressing *E. coli* colonies with 400 nm LED, 460 nm LED, 450-490 nm or 480-520 nm light from a 300 W xenon lamp does not generate any far-red fluorescence. We then put in other mutations of psmOrange relative to mOrange (i.e. S21T, F100Y, L125M, K163M; psmOrange numbering) into O-GECO-Q64L. The result variant, named as pcO-GECO0.2, displays significant improved folding and chromophore maturation. However no photoconversion was observed. As a control, modest photoconversion of psmOrange was observed using the same illumination and imaging conditions. Considering the low photocontrast and intensive light required to convert psmOrange, we did not pursue further directed evolution on it. Nevertheless psmOrange2, an improved psmOrange variant with higher photoconversion contrast and faster photoconversion kinetics, was recently reported [226]. Introducing the 6 additional mutations in psmOrange2 to pcO-GECO0.2 and further directed evolution might give a photoconvertible O-GECO.

4.2.4 Creation of circularly permuted variants of mMaple

We next turned to Kaede-type pcFPs for which there exist no GECO variants with high sequence similarity. Accordingly, we undertook a complete swap of the cpFP portion of G-GECO1.1 with a cp version of mMaple, a Kaede-type green-to-red pcFP described in Chapter 2 and Chapter 3. We first created a series of cpmMaple variants with the original termini connected by a (Gly)₂Ser(Gly)₂ linker and new termini at residues 140, 145, 193, and 207. Not unexpectedly, we found that variants with new termini far away from the chromophore (i.e. position 193 and 207) remained brightly fluorescent and photoconvertible (Figure 4.2 and Table 4.1). In contrast, cp variants with new termini near to the chromophore in β -strand 7 (i.e. position 140 and 145) were substantially dimmer though

still photoconvertible (Figure 4.2). In-colony brightness of cpmMaple207 is significantly dimmer than that of cpmMaple193. However, the purified proteins had similar brightness (Table 4.1), suggesting that it is the folding or maturation of the protein being impaired by circularly permutation.

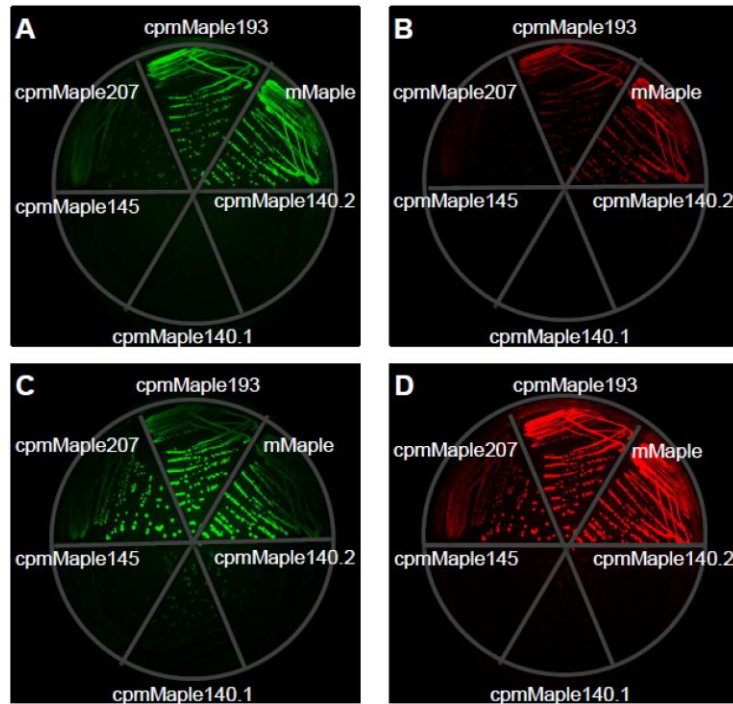


Figure 4.2 Fluorescent images of *E. coli* expressing cpmMaple variants on a Petri dish. (A-B) Pre- (A) and post-photoconversion (B) images of *E. coli* grown for 15 hours at 37°C. (C-D) Green (C) and red (D) fluorescence of the same plate after another 5-days at room temperature. Additional photoconversion was used to convert newly synthesized FPs to red before acquisition of (D). Substitution at the new N- and C- termini are His and Asp for cpmMaple193, Arg and stop codon for cpmMaple207, Gly and Thr for cpmMaple145, Thr and stop codon for cpmMaple140.1 and Asp and Ala for cpmMaple140.2.

Table 4.1 Properties of mMaple and its circularly permuted variants

State	Protein	λ_{exc} (nm)	λ_{em} (nm)	ϵ^a	Φ	Brightness ^b
green	mMaple	489	505	15	0.74	11
	cpmMaple193	489	505	18	0.80	14
	cpmMaple207	489	505	14	0.84	12
red	mMaple	566	583	30	0.56	17
	cpmMaple193	566	582	32	0.50	16
	cpmMaple207	566	582	29	0.53	15

^a Unit of $\text{mM}^{-1} \text{cm}^{-1}$. ^b Product of ϵ and Φ in $\text{mM}^{-1} \text{cm}^{-1}$.

4.2.5 Creation of GR-GECO0.1 and its optimization

Since the variant with new termini at position 145 was the closest analogue of the cpFP portion of other single FP-based Ca^{2+} indicators, we introduced it in place of the cpFP of G-GECO1.1, despite its low brightness. Satisfyingly, we found that the resulting construct was weakly fluorescent but capable of green-to-red photoconversion. It also displayed a small increase in red fluorescence upon Ca^{2+} binding (~20%). We named this first generation pcGECO as green-to-red (GR)-GECO0.1 and used it as the initial template for an extensive process of optimization for improved brightness and Ca^{2+} response in the red channel.

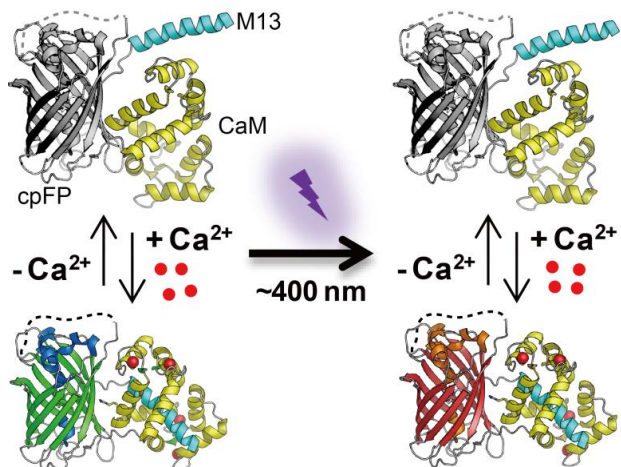


Figure 4.3 Schematic representation of a single FP-based green-to-red photoconvertible Ca^{2+} indicator. The indicator can be photoconverted from green fluorescent into red fluorescent with near-UV illumination. In the absence of Ca^{2+} , both the pre-photoconversion (green) and the post-photoconversion (red) are dimly fluorescent. Binding of Ca^{2+} leads to significant signal increase on fluorescence (either green or red). Photoconversion happens regardless of the presence of Ca^{2+} .

Initial efforts to optimize the response of GR-GECO focused on exploring different lengths and compositions for the peptide sequences that linked the cpFP domain to CaM and M13 (Table 4.2). This effort ultimately led to the identification of GR-GECO0.2 with two additional residues (Gly-Tyr) in the linker between M13 and the cpFP domain (Table C1 in Appendix C). We next turned our attention to the presumed interface between the cpFP domain and the CaM/M13 complex. X-ray crystallographic analysis of G-CaMP2, plus various mutagenesis studies have identified a number of residues (e.g. Glu61, Arg377, Asn380, and Asp381; G-CaMP2 numbering) that are in this interface and help to communicate the Ca^{2+} -dependent conformational change into a change of the FP chromophore environment [166, 167]. Libraries of GR-GECO variants, created by genetic randomization of these positions, were screened for improved efficiency of photoconversion and improved Ca^{2+} response for the red species. This effort led to the identification of GR-GECO0.3 with 3 additional substitutions (E61P/N380W/D381G) (Table

C1) and improved brightness, photoconversion, and response to Ca^{2+} (~100%) relative to GR-GECO0.2.

Table 4.2 Termini library of GR-GECOs

Termini library	N-terminus position	C-terminus position	Library size	Brightness	Ca^{2+} response
1	146 147	141 142	$64 \times 64 =$ 4096	---	NA
2	145	143 144	$32 \times 64 =$ 2048	+	+
3	144 (V145Y ^a)	143 144	$32 \times 64 =$ 2048	++	+
4	144 (V145Y ^b)	145	$32 \times 32 =$ 1024	++	++

^a A forward primer containing V145Y, a mutation decided from library 2, was used. ^b Mutation V145Y from now on will be named as +2Y while mutation E144G named as +1G since residues 144 and 145 also exists in the C terminus of cpmMaple.

Subsequently, we resorted to the use of directed evolution through iterative rounds of library creation by error-prone PCR and library screening for improved photoconversion efficiency and increased Ca^{2+} response of the red species. After 16 rounds, we arrived at 5 variants that showed satisfying photoconversion and Ca^{2+} response. Based on an assessment of both in vitro properties and performance in live cell imaging, we selected two variants, designated as GR-GECO1.1 and GR-GECO1.2 respectively, as the most preferred final versions. GR-GECO1.1 displays satisfying Ca^{2+} response at both green and red species, whereas GR-GECO1.2 exhibits excellent Ca^{2+} response at the red species (Figure 4.7 and Table 4.3). GR-GECO1.1 has 28 mutations or insertions distributed around the protein, while GR-GECO1.2 has 24 mutations or insertions, most of which similar to those in GR-GECO1.1 (Figure 4.4, Figure 4.5 and Table C1)

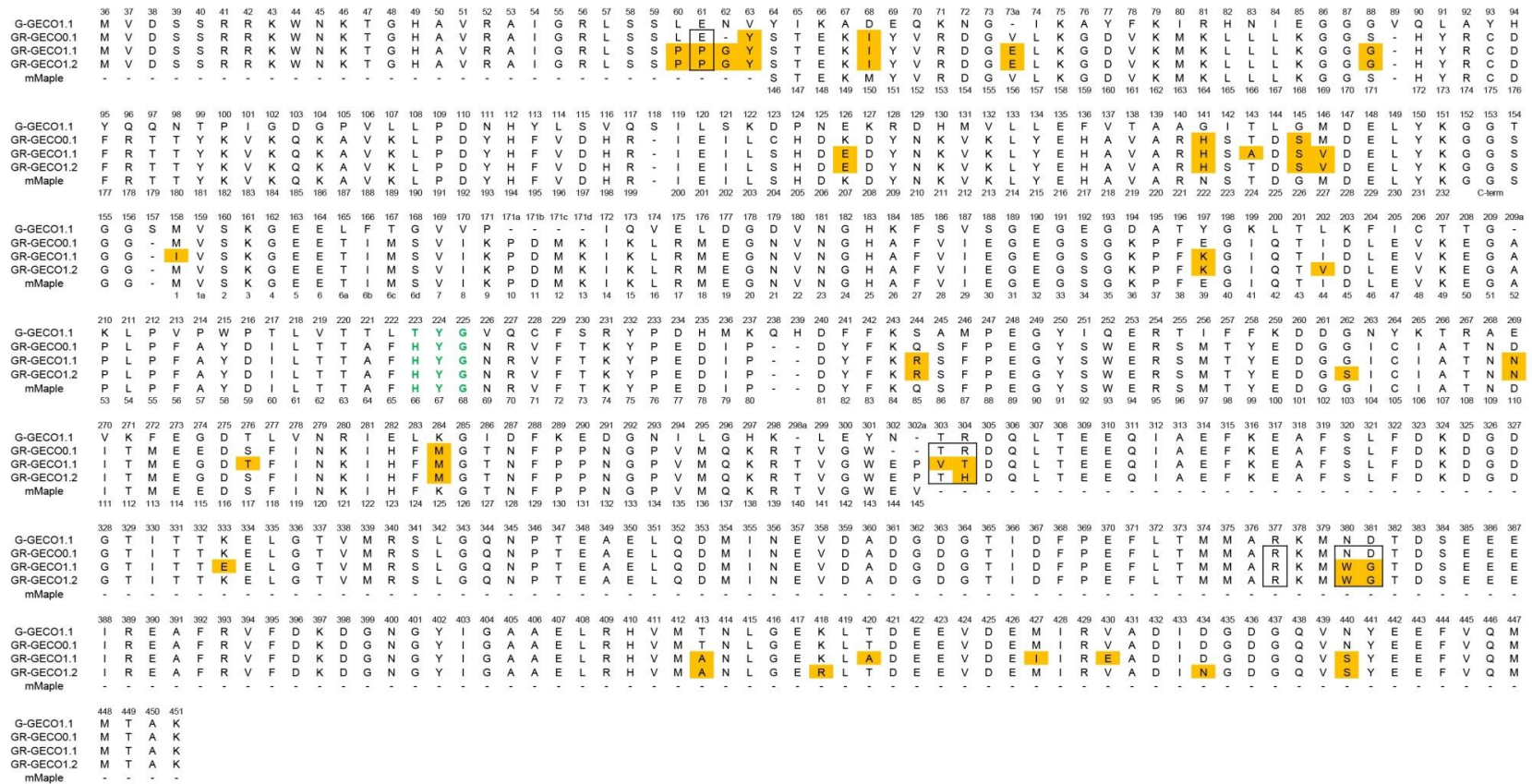


Figure 4.4 Sequence alignment of GR-GECOs, G-GECO1.1 and mMaple. Positions targeted for saturation or semi-saturation mutagenesis are highlighted with a box. Mutations found are highlighted with orange shade. Letters in green are the chromophore forming residues. Top row numbering corresponds to G-CaMP and G-GECO numbering convention. Bottom row numbering corresponds to mMaple numbering convention.

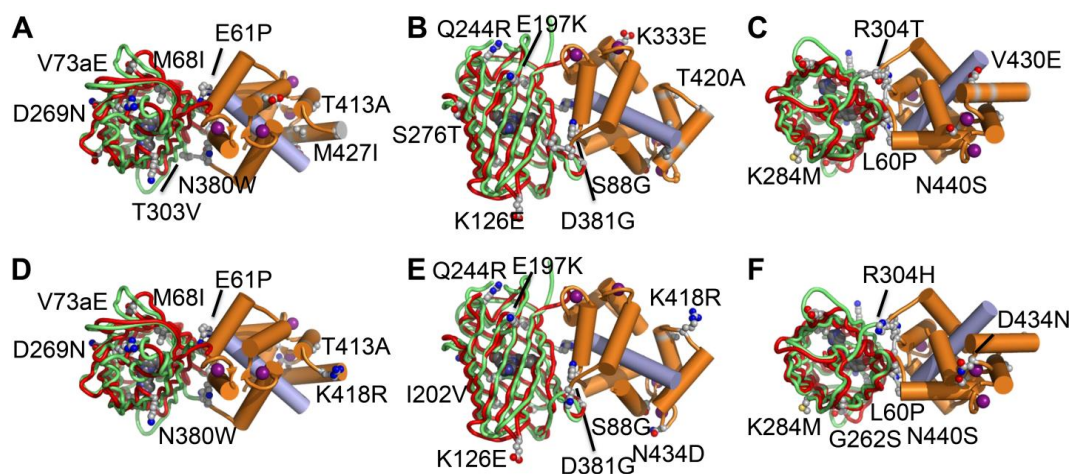


Figure 4.5 Location of substitutions in GR-GECOs. Structure of G-CaMP2 (PDB: 3EVR) [166] and the pre-photoconversion state of Eos (PDB: 1ZUX) [227], a photoconvertible FP which is a close homolog of mMaple, are aligned to represent the overall structure of GR-GECOs. Structure of cpGFP in G-CaMP2 is represented in green, while structure of Eos is represented in red. (A-C) Top (A), side (B) and bottom (C) views of GR-GECO1.1. Substitutions not shown in the picture includes 2+62G, 1+63Y, N141H, T143A, G145S, M146V and M158I. “+” means extra residues before cpmMaple (Figure 4.4). (D-F) Top (D), side (E) and bottom (F) views of GR-GECO1.2. Substitutions not shown in the picture includes 2+62G, 1+63Y, N141H, G145S and M146V (Figure 4.4).

4.2.6 *In vitro* characterization of GR-GECOs

The absorbance spectra of both GR-GECOs reveal that, in the absence of Ca^{2+} , both the green and red state chromophores exist predominantly in their protonated (non-fluorescent) forms at pH 7.4 (Figure 4.6). Binding of Ca^{2+} shifts the chromophore into the anionic fluorescent state. Accordingly, the GR-GECOs share a similar Ca^{2+} sensing mechanism to that of G-CaMPs and G-GECOs. As observed for the GECO variants, the Ca^{2+} sensing mechanism is best described as a Ca^{2+} -dependent shift in chromophore $\text{p}K_a$ to a lower value, with larger shifts giving rise to larger fluorescence responses.

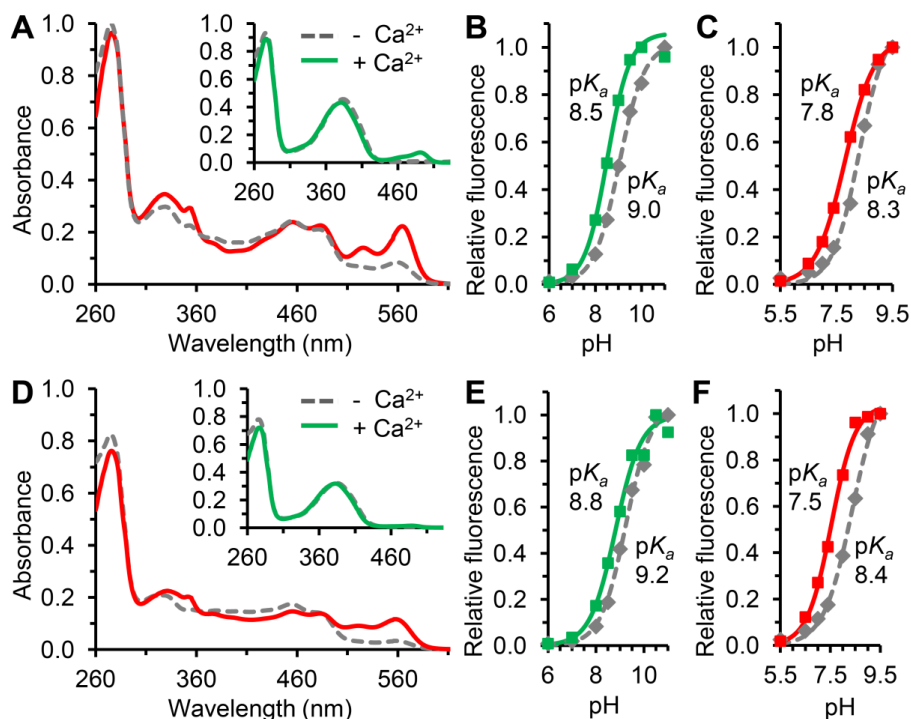


Figure 4.6 Absorbance spectra and pH titration of GR-GECOs. (A) Absorbance spectra of the red and green (inset) species of GR-GECO1.1 with (solid red or green) and without (dash grey) Ca²⁺. (B-C) pH titration of the green (B) and the red (C) species of GR-GECO1.1 with (green or red square) or without (grey diamond) Ca²⁺. (D-F) Absorbance spectra and pH titration of GR-GECO1.2. Line color and line style representations are same to those in (A-C).

Spectroscopic characterization of purified protein revealed that the green and red states of GR-GECO1.1 have a 8.6-fold and a 6.1-fold increase in fluorescence upon binding Ca²⁺, respectively (Figure 4.7A, B and Table 4.3). GR-GECO1.2 exhibits 3.9-fold and 6.7-fold increases, respectively (Figure 4.7C, D and Table 4.3). Since reversible photoswitching between the protonated and deprotonated states of red species has been observed with mEos2 and mMaple (see Chapter 3) [209], we examined GR-GECOs for such behavior. Fortunately, prolonged illumination of the protonated state ($abs_{max}=455$ nm) does not reversibly shift the equilibrium towards the deprotonated state, either in the presence or absence of Ca²⁺ (Figure 4.8). However, we did notice that excitation of the protonated state of the green species ($abs_{max}=383$ nm) of GR-GECO1.2 resulted in significant green

emission (Figure 4.7C). In contrast, mMaple has negligible green fluorescence when excited at this absorbance peak (Figure 2.12). We reasoned that this green emission is likely the result of excited state proton transfer (ESPT) [171], and this process must be competing with the photoconversion process. In agreement with this hypothesis, we observed slower photoconversion in GR-GECO1.2 compared with GR-GECO1.1 (Figure 4.9). Notably, the photoconversion rate of GR-GECO1.1 in the presence of Ca^{2+} is significantly faster than the rate without presence of Ca^{2+} . This discrepancy is potentially useful for generation of a Ca^{2+} integrator that would allow researchers to integrate Ca^{2+} flux over a period of time and study the signal accumulation during physiological processes such as learning and memory.

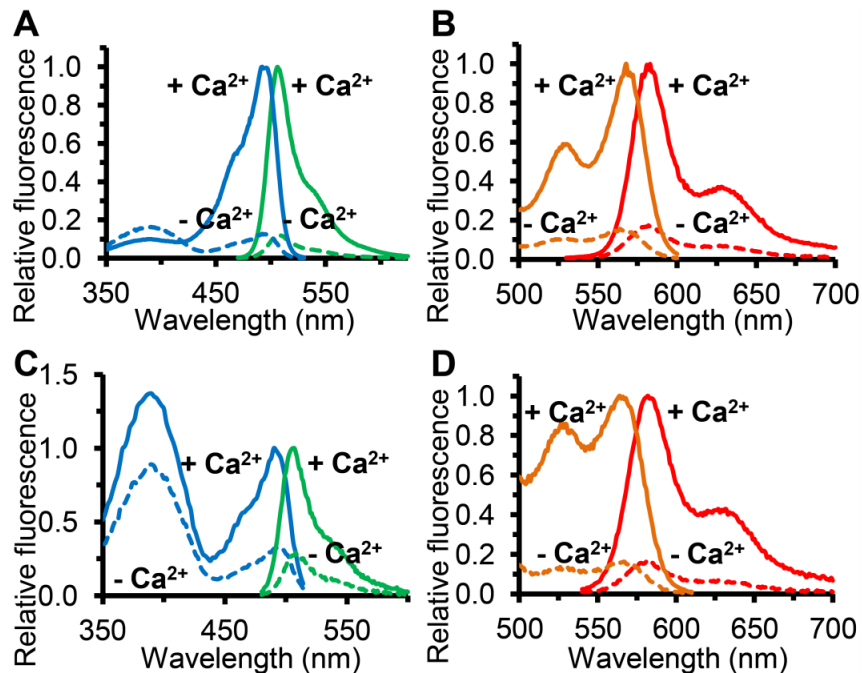


Figure 4.7 Excitation and emission spectra of GR-GECOs. (A) Excitation (blue) and emission (green) spectra of the green species of GR-GECO1.1, either with (solid) and without (dash) Ca^{2+} . (B) Excitation (orange) and emission (red) spectra of the red species of GR-GECO1.1, either with (solid) and without (dash) Ca^{2+} . (C-D) Excitation and emission spectra of green (C) and red (D) species of GR-GECOs. Line color and line style representations are same to those in (A) and (B).

Table 4.3 Properties of GR-GECOs

Protein	Ca ²⁺	λ_{abs} (nm) with ϵ in parenthesis		λ_{em} (nm)	Φ	B ^a	pK _a	Intensity change with Ca ²⁺	$(F_{\text{Ca}} - F_{\text{EGTA}})/F_{\text{EGTA}}$ in HeLa with number of trials in parenthesis ^b
GR- GECO0.5 (green)	-	386 (31)	490 (0.4)	509	0.77	0.3	9.4	2.9	2.7 ± 0.4 (15)
	+	383 (27)	490 (1.4)	507	0.63	0.9	8.8		
GR- GECO1.1 (green)	-	383 (29)	487 (1.5)	508	0.70	1.1	9.0	3.4	9.4 ± 1.7 (18)
	+	381 (27)	491 (5.7)	506	0.62	3.5	8.5		
GR- GECO1.2 (green)	-	384 (30)	488 (1.1)	506	0.65	0.7	9.2	2.2	6.2 ± 1.3 (17)
	+	383 (29)	488 (2.2)	506	0.71	1.6	8.8		
GR- GECO0.5 (red)	-	558 (4.6)		582	0.51	2.3	8.2	2.8	2.7 ± 0.2 (15)
	+	557 (12.7)		584	0.51	6.5	7.5		
GR- GECO1.1 (red)	-	559 (5.5)		582	0.46	2.5	8.3	3.2	3.4 ± 0.6 (24)
	+	564 (14.1)		583	0.57	8.0	7.8		
GR- GECO1.2 (red)	-	558 (3.0)		581	0.39	1.2	8.4	4.6	5.1 ± 1.3 (51)
	+	558 (10.6)		582	0.51	5.4	7.5		

^a Brightness is defined as the product of ϵ (mM⁻¹·cm⁻¹) and Φ . ^b Cells were treated first with histamine, then with EGTA/ionomycin to give a minimum fluorescence F_{EGTA} , and finally with Ca²⁺/ionomycin to give a maximum signal F_{Ca} . Each trial represents an individual transfected cell on which systematic calibration experiments were performed.

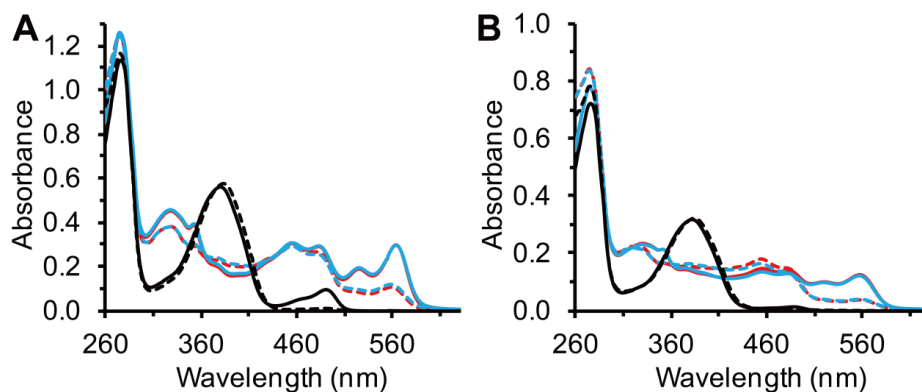


Figure 4.8 Effect of 460 nm illumination on photoswitching. Absorbance spectra of GR-GECO1.1 (A) and GR-GECO1.2 (B) before photoconversion (black), after 30 min illumination with 400 nm LED (red) and after another 30 min illumination with 460 nm LED (blue). The Ca^{2+} free states are represented by dashed lines, while the Ca^{2+} -bound states are represented by solid lines. Unlike mMaple, the red species of GR-GECOs do not undergo observable photoswitching with prolonged illumination at 460 nm.

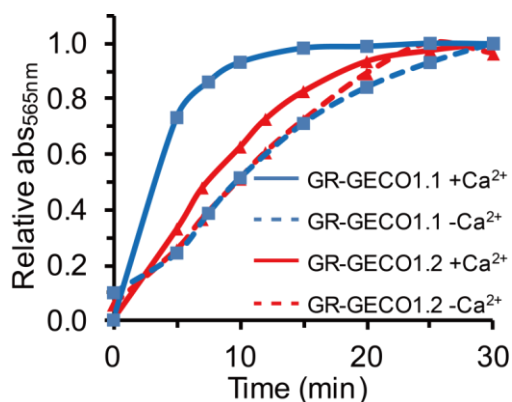


Figure 4.9 Photoconversion profile of GR-GECOs. The 405 nm LED illumination system was used for photoconversion. Photoconversion of GR-GECO1.1, as monitored as the generation of the red anionic state (i.e. absorbance at 565 nm), is faster than that of GR-GECO1.2.

In vitro Ca^{2+} titrations enabled us to determine the K_d (defined as the concentration of Ca^{2+} that gives half the maximum fluorescence increase) for GR-GECO1.1 (86 nM and 54 nM for the green and red species, respectively) and GR-GECO1.2 (74 nM and 90 nM, respectively) (Figure 4.10 and Table 4.4). Further characterization by stopped-flow fluorescence spectroscopy demonstrated that the relaxation rate profiles of GR-

GECCO1.1 at different Ca²⁺ concentration followed monoexponential increase functions and give Hill coefficients and K_d values consistent to what were obtained by Ca²⁺ titration (Figure 4.11A, B and Table 4.4). In contrast, GR-GECCO1.2 exhibited complex stopped-flow kinetics and the data could not be satisfactorily fit to a simple kinetic model (Figure 4.11C, D).

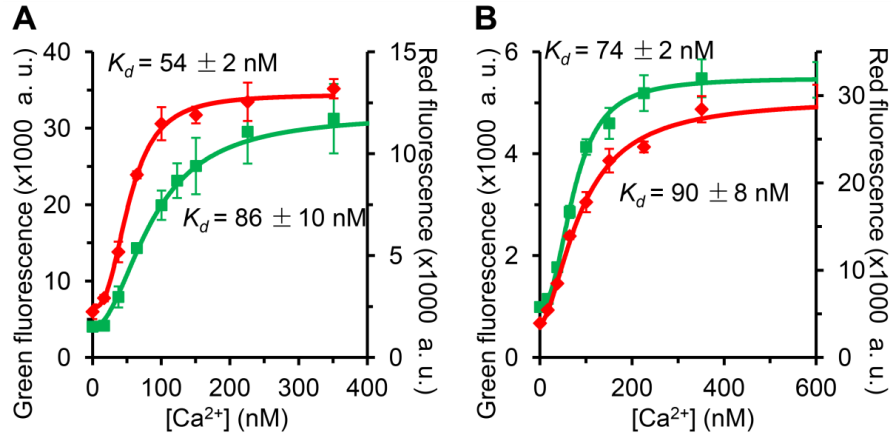


Figure 4.10 Ca²⁺ titration of GR-GECCOs. (A-B) Fluorescence intensity plotted as a function of the [Ca²⁺] for the green (green square) and red species (red diamond) of GR-GECCO1.1 (A) and GR-GECCO1.2 (B). K_d and Hill coefficients are determined by fitting to the equation $F_i = F_0 + (F_{max} - F_0) \times \frac{[Ca^{2+}]^n}{(K_d^n + [Ca^{2+}]^n)}$.

Table 4.4 Ca²⁺-binding equilibrium of GR-GECCOs

Protein	State	k_{on} (M ⁻ⁿ s ⁻¹)	k_{off} (s ⁻¹)	$n_{kinetic}$	$K_{d,kinetic}$ (nM) ^a	$K_{d,static}$ (nM)	n_{static}
GR-GECCO1.1	green	1.50×10^{-5}	0.536	1.98	200	86	2.0
	red	4.90×10^{-7}	0.582	2.56	235	54	2.6
GR-GECCO1.2	green	ND	ND	ND	ND	74	1.7
	red	ND	ND	ND	ND	90	1.7

^a $K_{d,kinetic} = (k_{off}/k_{on})^{1/n}$

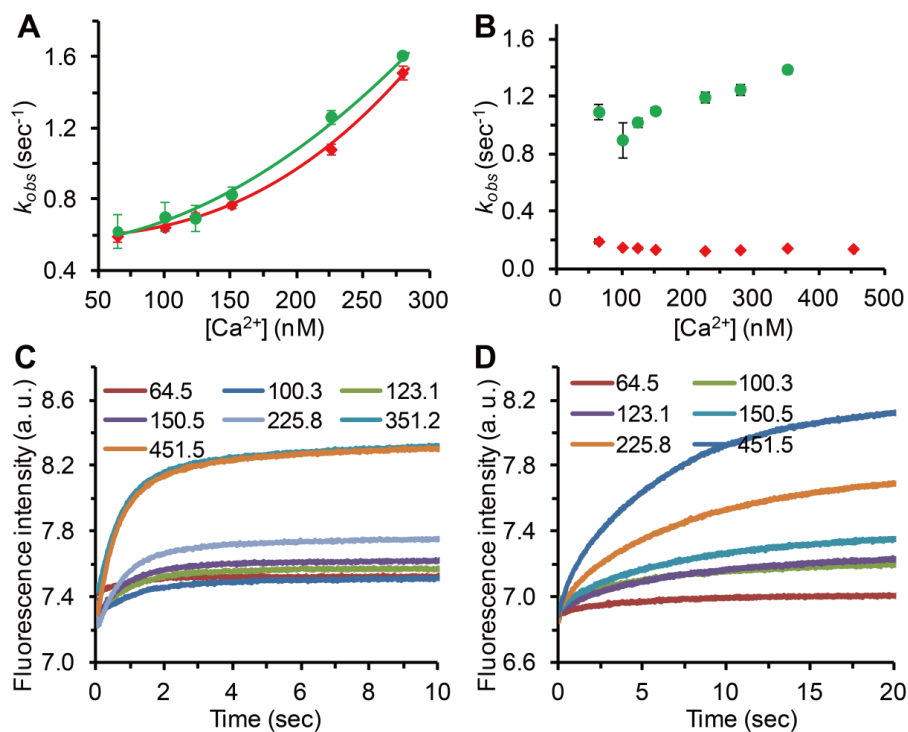


Figure 4.11 Stopped-flow kinetic characterization of GR-GECOs. (A-B) Observed relaxation rate constants (k_{obs}) plotted as a function of $[Ca^{2+}]$ for both green (green circle) and red species (red diamond) of GR-GECO1.1 (A) and GR-GECO1.2 (B). Association (k_{on}) and dissociation (k_{off}) rate constants of GR-GECO1.1 were determined by fitting to the equation $k_{obs} = k_{on}[Ca^{2+}]^n + k_{off}$. Values are tabulated in Table 4.4. Data for GR-GECO1.2 could not be satisfactorily fit using the same procedure, suggesting that the binding kinetics are more complex for this variant. (C-D) Relaxation trace of the green species (C) and the red species (D) of GR-GECO1.2. Numbers in the legend are the $[Ca^{2+}]$ in nM.

4.2.7 *In vivo* characterization of GR-GECOs

To determine how GR-GECOs perform in live cell imaging, we expressed them in HeLa cells and, following partial photoconversion by illumination of the whole field of view with violet light, used established protocols to determine their response to changes in cytoplasmic Ca^{2+} concentrations [216]. Specifically, GR-GECOs were expressed in the cytoplasm of HeLa cells. In a time-lapse imaging experiment, histamine was added into the cell culture to induce $[Ca^{2+}]$ fluctuation in cytoplasm. The cytoplasmic Ca^{2+} was then depleted by addition of EGTA/ionomycin to find out the fluorescence of the Ca^{2+} indicators in the apo-state. In the

last step, CaCl_2 /ionomycin was applied to find out the maximum signal when the indicators are saturated. As expected, we observed oscillations, decreases to minimum and increases to maximum in the fluorescence intensity when cells were treated with histamine, EGTA/ionomycin and Ca^{2+} /ionomycin, respectively (Figure 4.12). In cells, the dynamic range for the green and red species, respectively, was 9.4-fold and 3.4-fold for GR-GECO1.1, and 6.2-fold and 5.1-fold for GR-GECO1.2 (Figure 4.12 and Table 4.3). Observation of increased fluorescence change upon addition of ionomycin/ Ca^{2+} is due to transient pH elevation in the cytoplasm caused by ionomycin and high Ca^{2+} . It was reported that cytosolic pH rose to about 8.5 upon addition of ionomycin/20 mM Ca^{2+} , while remained neutral during histamine and ionomycin [228].

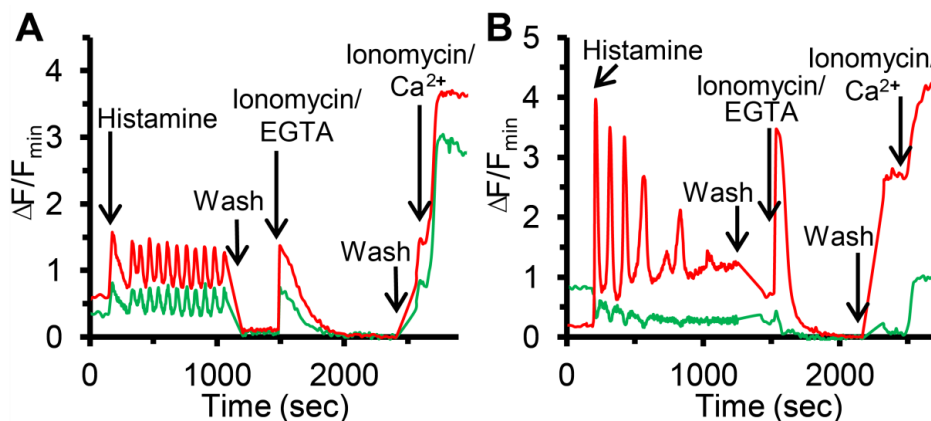


Figure 4.12 Representative timecourse traces for HeLa cells expressing cytoplasmic GR-GECO1.1 (A) or GR-GECO1.2 (B). For GR-GECO1.2, the red species displays a larger dynamic range and greater photostability than the green species. Addition of histamine induces the oscillation of cytoplasmic $[\text{Ca}^{2+}]$, while addition of ionomycin/EGTA cause a rapid increase of Ca^{2+} influx followed by the cytoplasmic Ca^{2+} being chelated by EGTA. In the last step, addition of ionomycin/ CaCl_2 causes a rapid increase of $[\text{Ca}^{2+}]_{\text{cyto}}$ and saturates the Ca^{2+} indicators.

To demonstrate the primary advantage of a highlightable Ca^{2+} indicator, we used light from a 405 nm laser to highlight single cells in a population of GR-GECO1.1 expressing HeLa cells (Figure 4.13). As expected, only the highlighted cells were observed in the red emission channel, and they

exhibited typical histamine-induced Ca^{2+} oscillations. Similarly, our collaborator Dr. Tomoki Matsuda from Osaka University found that single cells in a population of primary neurons transfected with GR-GECO1.2 could be highlighted and spontaneous Ca^{2+} oscillations could be imaged in the red emission channel (Figure C1 in Appendix C). We anticipate that highlightable Ca^{2+} indicators will facilitate the dissection and study of the Ca^{2+} signaling in neuronal networks.

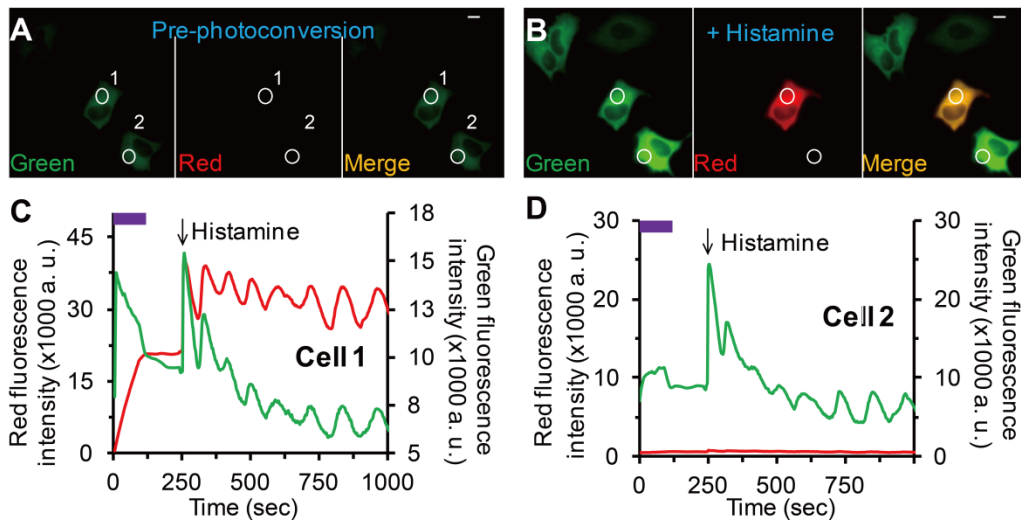


Figure 4.13 Imaging histamine-induced Ca^{2+} oscillations in highlighted HeLa cells. (A) Green (left), red (middle) and merged (right) fluorescence images of HeLa cells expressing GR-GECO1.1 prior to photoconversion. (B) Same field of view as (A) after photoconversion of cell 1 and addition of histamine. Scale bar 10 μm . (C) Fluorescence intensity plotted as a function of time for cells 1 and 2. Purple bar indicates the duration of photoconversion.

4.3 Conclusions

Two of the most powerful implementations of fluorescent protein (FP) technology are the “highlighters” that can be converted from non-fluorescent to fluorescent or from one color to another by illumination, and the Ca^{2+} indicators. We have combined the properties of both of these FP classes into a single construct to produce a highlightable Ca^{2+} indicator that will enable researchers to spectrally mark a single cell in a transfected tissue and image its intracellular Ca^{2+} dynamics. As described in this

chapter, we explored four different protein design strategies to create such a hybrid tool. We attempted to convert existing green-emissive, red-emissive and orange-emissive GECOs into photomodulable (referred to photoactivable and photoconvertible here) by introducing mutations that are critical for converting the corresponding FPs from non-photomodulable into photomodulable. However, little success was achieved with these approaches. The strategy that ultimately proved successful involved the creation of a circularly permuted version of a green-to-red photoconvertible FP and its introduction into a G-CaMP-type single FP-based Ca^{2+} indicator. Optimization by directed evolution led to the identification of two promising variants that exhibit excellent photoconversion properties and have an up to 4.6-fold increase in red fluorescence intensity upon binding Ca^{2+} . The utility of these variants was demonstrated in HeLa cells, as well as neuron cells. By enabling the selective highlighting and imaging of individual cells in densely transfected tissues, these indicators will enable researchers to explore intra- and intercellular signaling with unprecedented experimental control and versatility.

4.4 Materials and methods

4.4.1 General methods and materials

See Section 2.4.1.

4.4.2 Generation of circularly permuted variants of mMaple

Overlap extension PCR was used to generate circularly permuted variants of mMaple. For example, to generate the cpmMaple193 variant, cDNA encoding residues 193-232 was amplified using a 5' primer with a *XhoI* site and a 3' primer with an overhang encoding "GGSGG". In parallel, the cDNA encoding residues 1-193 was amplified using a 5' primer with a complementary sequence encoding "GGSGG" and a 3' primer with two

extra Gly residues followed by an *EcoRI* site. The terminal residues (i.e. positions 193 and 192) were randomized with an NNK sequence that encodes for all 20 amino acids using 32 codons. The above two fragments were then stitched together by overlap extension PCR to generate cpmMaple193. CpmMaple207, 145, 140, 143 and 144 were constructed by similar methods, except that the last three constructs had only one extra glycine residue at the new C-terminus. The assembled cDNA was digested with *XhoI* and *EcoRI* and ligated into similarly digested pBAD/His B (Invitrogen). Electrocompetent *Escherichia coli* strain DH10B (Invitrogen) was transformed by electroporation and plated on LB (Luria–Bertani)/agar plates supplemented with ampicillin (0.1 mg/ml) and L-arabinose (0.02%). Plates were incubated overnight at 37°C and then at room temperature prior to screening.

4.4.3 Plasmid library creation and site-directed mutagenesis

Linker libraries of GR-GECOs were generated by amplifying the cDNA of cpmMaple145 with various 5' primers with a *XhoI* site and 3' primers with a *MluI* site. These primers encode for polypeptide sequence that starts or stops at various positions between residues 141 to 147. The purified PCR products were digested with *XhoI* and *MluI* and ligated into similarly digested pTorPE. The pTorPE is a modified pBAD/His B contains a TorA periplasmic export sequence ahead of the 6×His tag [168].

Randomly mutated libraries were created by error-prone PCR [182], using a 5' primer with an *XbaI* site annealing to the 6×His tag and a 3' primer with a *HindIII* site. In addition, a modified staggered extension process (StEP) PCR was used to create libraries that contained various combinations of beneficial mutations found during previous rounds of screening [183]. Briefly, a total of 0.1 pmol pBAD/His B plasmids, containing the genes for 5-15 different GR-GECO variants selected from randomly mutated libraries, were used as the template mix. The mixture was PCR amplified for 80-100 short cycles (94°C for 20 s, 55°C for 15 s

and then 72°C for 5 s), followed by a final extension step at 72°C for 5 min.

Saturation or semi-saturation mutagenesis at specific residues was performed using overlap extension PCR or QuickChange Site-Directed Mutagenesis Kit (Agilent Technologies). Regardless of library assembly method (i.e. error-prone PCR, saturation mutagenesis, or StEP PCR), gene libraries were digested with *XbaI* and *HindIII* and inserted into a similarly digested pTorPE. Electrocompetent *E. coli* strain DH10B (Invitrogen) was transformed and plated on LB (Luria–Bertani)/agar plates supplemented with ampicillin (0.1 mg/ml) and L-arabinose (0.0004-0.0008%). Plates were incubated overnight at 37°C and another 10-24 hours at room temperature prior to screening.

4.4.4 Plasmid library screening

Since the expressed protein was exported to the high Ca^{2+} environment of the periplasm, colonies containing functioning Ca^{2+} sensors would, in principle, manifest brighter fluorescence. To screen for both photoconversion (or photoactivation) and Ca^{2+} response, we performed a two-step screening procedure. In the first step, plates containing colonies were photoactivated or photoconverted for 5-15 min using a custom built illumination source composed of six 9×11 arrays of 405 nm light emitting diodes (LED) (OptoDiode Corporation, Newbury Park, CA). Colonies exhibiting the brightest fluorescence in the post-photoactivation/photoconversion channel were selected. For directed evolution of GR-GECOs, the bright red fluorescence could be result of either high photoconversion or high Ca^{2+} response, or both. To ensure the selected variants had satisfactory Ca^{2+} response, they were grown at 37°C overnight in LB media supplemented with ampicillin and L-arabinose, then transferred to 30°C for another 10-24 hr. Cells were pelleted by centrifuge, washed with a MgSO_4 /EGTA solution to remove excess Ca^{2+} , and lysed in B-Per (Pierce). A portion of this B-Per lysate was taken out for

spectrofluorometric testing of the Ca^{2+} response of the green species. The remainder of the lysate was photoconverted using the aforementioned chamber. Both the pre- or post- photoconversion proteins were diluted into KCl-MOPS buffer (30 mM MOPS, 100 mM KCl, pH 7.4) in a 96-well plate for Ca^{2+} response test. Three fluorescence intensities was recorded for each sample: the initial value (F_{ini}), the value after addition of 0.4 mM (final concentration) EGTA (F_{EGTA}), and the value after addition of equivalent amount of Ca^{2+} ($F_{\text{Ca}^{2+}}$). Variants showing the largest $(F_{\text{Ca}^{2+}} - F_{EGTA})/F_{EGTA}$ (i.e. $\Delta F/F_{min}$) value were considered as winners. Plasmids for these variants were extracted for sequencing and served as templates for subsequent library generation.

4.4.5 Protein purification and characterization

Protein purification was carried out as previously described in Section 2.4.4 except that a lower concentration of arabinose (0.002%) was used to induce the expression of GR-GECOs. The buffer was changed into PBS (pH 7.4) for cpmMaple variants or KCl-MOPS buffer (30 mM MOPS, 100 mM KCl, pH7.4) for GR-GECOs. All absorption measurements were acquired on a DU-800 UV–visible spectrophotometer (Beckman). Unless otherwise indicated, all fluorescence spectra were recorded on a QuantaMaster spectrofluorometer (Photon Technology International) and have been corrected for the detector response. For protein samples representative of the apo state (i.e. Ca^{2+} free state), concentrated protein sample was diluted into Ca^{2+} free KCl-MOPS buffer (30 mM MOPS, 100 mM KCl and 4 mM EGTA, pH7.4). For protein samples representative of the saturated state (i.e. Ca^{2+} bound sate), concentrated protein was diluted into Ca^{2+} containing KCl-MOPS buffer (30 mM MOPS, 100 mM KCl and 4 mM CaCl_2 , pH7.4). To obtain the red state of the GR-GECOs, concentrated protein was subjected to illumination with the custom built 405 nm light source for 10 min.

Molar extinction coefficients (ϵ) and quantum yields (Φ) determination. Molar extinction coefficients were determined as described in Section 2.4.4. To determine the fluorescence quantum yields, purified proteins were diluted into citrate-borate buffer (pH 9.0) with 4 mM EGTA or with 4 mM Ca^{2+} . The high pH buffer was used in order to yield enough absorbance and emission of the deprotonated state at low protein concentration. Fluorescein in 10 mM NaOH ($\Phi = 0.95$) [190], and the red species of mMaple in PBS (pH 7.4) ($\Phi = 0.53$), were used as Φ standards for the green and red states, respectively.

pH titration for pK_a determination. For determination of the pH dependence, concentrated GR-GECO proteins were diluted 1:50 into a series of pH-adjusted borate/citrate buffers with 4 mM EGTA or 4 mM CaCl_2 in a 96-well black clear bottom plate (Corning). Fluorescence was measured using a Safire2 plate reader.

Ca^{2+} titration for K_d determination. To determine the K_d of GR-GECOs for binding to Ca^{2+} , a series of KCl-MOPS buffers containing different amount of Ca^{2+} were prepared following the instructions provided in the Calcium Calibration Buffer Kit (Invitrogen). Purified GR-GECOs was then diluted 1:100 into the Ca^{2+} buffer and their fluorescence measured with a Safire2 plate reader (Tecan). Data was fit with a Hill equation: $F_i = F_0 + (F_{max} - F_0) \times \frac{[\text{Ca}^{2+}]^n}{(K_d^n + [\text{Ca}^{2+}]^n)}$, where n is the Hill coefficient and F_0 and F_{max} are the fluorescence without Ca^{2+} and with saturated Ca^{2+} , respectively. The calculated K_d represents the concentration of Ca^{2+} when the fluorescence change of the indicator is half of its maximum value.

Stopped-flow spectroscopy for Ca^{2+} -binding kinetics determination. A SX20 stopped-flow spectrometer (Applied Photophysics) equipped with a 150 W xenon lamp and a fluorescence photomultiplier was used for stopped flow kinetic measurements. The GR-GECO indicator (in 30 mM MOPS, 1 mM EGTA and 100 mM KCl) was rapidly mixed (1:1) with a series of Ca^{2+} buffers. Excitation wavelength was set to 480 nm for the

green and 525 nm for red species, each with a slit width of 14 nm. A 535/50 nm filter (for green emission) and a 585/40 nm filter (for red emission) were used. The change in the fluorescence signal during rapid mixing provided the relaxation rate constants (k_{obs}) for the Ca^{2+} association reaction at various Ca^{2+} concentrations (from 60 nM to 500 nM). By fitting the observed data to the equation $k_{obs} = k_{on} \times [Ca^{2+}]^n + k_{off}$, the kinetic constants of Ca^{2+} association k_{on} and dissociation k_{off} were determined.

Photoconversion with and without Ca^{2+} . Purified protein was diluted into KCl-MOPS buffer with 4 mM EGTA or with 4 mM Ca^{2+} and subsequently photoconverted into the red state using the 405 nm LED. Absorbance spectra at different photoconversion times were recorded. Notably, prolonged photoconversion of GR-GECOs depletes their Ca^{2+} -responding capability, possibly due to photoinduced or singlet-oxygen dependent modifications of the protein residues.

4.4.6 Live cell imaging

Construction of mammalian expression vectors. To create the GR-GECO expression vectors, the gene was amplified with a 5' primer with a *BamHI* site and a 3' primer with an *EcoRI* site. The purified PCR products were digested and ligated into a similarly digested modified pcDNA3 plasmid [168]. Briefly, this modified pcDNA3 plasmid was made by deleting 2224 nucleotides (including the SV40 promoter, the SV40 origin of replication, the Neomycin ORF, and the SV40 poly A region) from the original 5.4 kb pcDNA3 (Invitrogen).

Maintenance of HeLa cells and transient transfection were performed as described in Section 3.3.3. Imaging was done 36–48 h after transfection. The culture medium was changed to HEPES (25 mM) buffered Hanks' Balanced Salt Solution (HHBSS) before imaging. An inverted Nikon Eclipse Ti microscope equipped with a 75 W mercury-xenon lamp (Hamamatsu), 20× and 40× objectives (Nikon), and a 16-bit QuantEM

512SC electron-multiplying CCD (Photometrics) was used to perform standard Ca^{2+} calibration experiment. Photoconversion was done using the epi-fluorescence illumination source and a DAPI excitation filter set (335-385 nm with a 400 nm dichroic). The following filter sets were used for imaging: 483-493 nm (excitation), 500-550 nm (emission), and 488 nm (dichroic) for the green channel; 538-548 nm (excitation), 563-637 nm (emission), and 543 nm (dichroic) for the red channel.

For the single cell highlighting and Ca^{2+} imaging experiment, a Nikon Eclipse equipped with same camera and lamp, a 60 \times oil-immersion Apo TIRF objective (NA 1.49) (Nikon), and a 408 nm photoactivation laser with a power of > 500 mW (Melles Griot) was used. The photoactivation laser delivers a point illumination with an approximate diameter of 40 pixels at the center of the field of view. The following filters were used: 465-495 nm (excitation), 515-555 nm (emission), and 505 nm (dichroic) for the green channel; 528-553 nm (excitation), 590-650 nm (emission), and 565 nm (dichroic) for the red channel. Image procession was done using ImageJ [194].

**Chapter 5 Engineering a monomeric version of yellow
fluorescent protein zFP538**

5.1 Introduction

The toolbox of FPs is rapidly expanding with new members being discovered in nature, as well as variants being engineered in the laboratory. Accordingly, the research community now has FPs with emission maxima spanning the visible spectrum. However, if one takes a closer look and considers FPs that are practically useful for imaging, one will find that there is a 30 nm-spectral-gap between the yellow FPs and the orange FPs. The most useful yellow FPs are Venus and Citrine, both obtained by substituting threonine at position 203 of avGFP variants with tyrosine [229, 230]. This substitution brings in a π - π stacking interaction between the aromatic side chain of Tyr203 and the phenolic ring in the chromophore (Figure 5.1B), which results in a 20 nm red-shift in the excitation and emission spectra [52] relative to wild-type avGFP. The emission maxima of Venus and Citrine are 528 nm and 529 nm, respectively. Although the engineered FPs mHoneydew and mBanana are yellow fluorescing (emission maxima 537 nm and 553 nm, respectively), their low brightness (6% and 12% of the brightness of EGFP, respectively) has limited them from widespread application. Hence, the next “redder” FPs on the list of practically useful FPs are mOrange2 ($\lambda_{em} = 565$ nm) and mKO ($\lambda_{em} = 559$ nm) [68, 84], leaving a ~30 nm spectra gap. Indeed, there is also a small number of naturally-occurring yellow-hue FPs (emission between 525 nm and 555 nm). These include phiYFP ($\lambda_{ex} = 525$ nm; $\lambda_{em} = 537$ nm) cloned from hydromedusa *Phialidium* sp. [46]; oleYFP ($\lambda_{ex} = 514$; $\lambda_{em} = 528$ nm) from hydromedusa *Obelia* sp. [231]; and zFP538 ($\lambda_{ex} = 528$; $\lambda_{em} = 538$ nm) from button polyp *Zoanthus* sp. [42]. However, these natural FPs all exist as oligomers [42, 46, 231], whereas it is widely accepted in the field of FP engineering that a monomeric quaternary structure is required for a FP to be practically useful as a fluorescence tag or FRET partner.

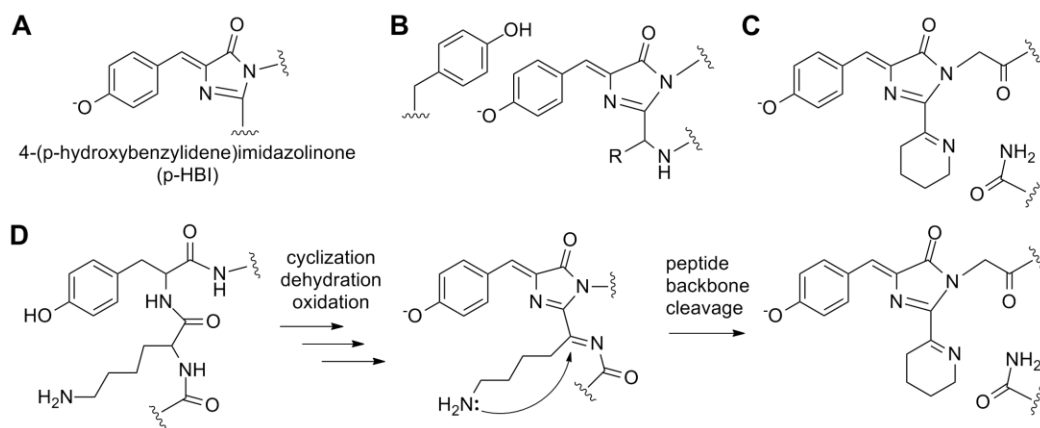


Figure 5.1 Chromophore structures and formation in yellow FPs. (A) 4-(*p*-hydroxybenzylidene)imidazolinone (p-HBI), the green fluorescing chromophore structure in avGFP. (B) Chromophore structure in some yellow FPs. An extra π - π stacking between Tyr203 and the phenolic ring of the p-HBI chromophore results in ~ 20 nm red-shift of the spectra [52]. The first amino acid residue in the triad could be Gly, Thr or Cys and is represented by “R” in the figure. (C) The three-ring chromophore structure in zFP538, where the imidazolinone ring has a cyclic-imine substitution [74]. (D) Proposed chromophore formation mechanism for zFP538. The N_{ξ} of Lys66 replaces the acylimine $C_{\alpha}=N$ bond at position 66, thereby forming a six-member hetrocyclic ring and cleaving the peptide backbone [74].

Remarkably, both oleYFP and phiYFP have the same structural solution for achieving yellow emission as the laboratory developed YFPs. That is, they have a typical 4-(*p*-hydroxybenzylidene)imidazolinone (p-HBI) chromophore (Figure 5.1A), and the bathochromic effect is the result of π - π stacking between a residue that is structural analogous to Tyr203 and the phenolate moiety of the chromophore (Figure 5.1B). The further red-shift in phiYFP (538 nm compared to 528 nm in oleYFP) is proposed to be due to additional hydrogen bonding interactions between the β -barrel and the chromophore [232].

In contrast to phiYFP and oleYFP, zFP538 possess a unique 3-ring chromophore (Figure 5.1C) formed from the amino acid triad Lys66-Tyr67-Gly68 as determined by a combination of mass spectrometry and X-ray crystallographic studies [74, 233]. Following the formation of a typical acylimine substituted p-HBI chromophore, the N_{ξ} of Lys66 undergoes

transimination to replace the acylimine $C_{\alpha}=N$ bond at position 66, leading to the formation of a six-member hetrocyclic ring and the breakage of the peptide backbone (Figure 5.1D) [74]. Although a Lys-Tyr-Gly-derived chromophore is also found in cyan FP amFP486, its mature chromophore structure resembles the typical p-HBI and lacks the third ring [42]. Therefore, the extra cyclization step in zFP538 must require an additional oxidation step and, possibly, assistance from surrounding residues in the chromophore environment.

The distinct 538 nm emission maximum of zFP538 and phiYFP fills the spectral gap between the greenish-yellow FPs and the orange FPs and is well separated from those of green and red FPs. This true yellow emission makes them excellent candidates for a third color choice in multi-color imaging. Moreover, they are excellent candidates for FRET acceptors when paired with cyan FPs and FRET donors when paired with red FPs. The genes for both zFP538 and phiYFP are commercially available, however as mentioned earlier, their application has largely been limited to use as gene expression reporters due to their notorious tendency to oligomerize. PhiYFP is a weak dimer and its improved variant TurboYFP, which is brighter and faster-maturing, remains a dimer as described on the commercial vendor's website (<http://www.evrogen.com>). Although a non-aggregating tetrameric zFP538 mutant has previously been reported [234], there have been no reports of a truly monomeric variant. In this chapter, I will describe our successful efforts to engineer and characterize the first monomeric variants of zFP538.

5.2 Results and discussion

5.2.1 Attempted engineering of monomeric yellow FP with a KYG chromophore through chromophore transplantation

In an attempt to generate a monomeric FP with the unique 3-ring chromophore structure of zFP538, we first attempted chromophore

transplantation, where the chromophore-forming residues of zFP538, and additional residues in close proximity, were transplanted into the structurally homologous and well characterized monomeric FP, mTFP1. Precedent for this approach came from the fact that modification of the chromophore-forming amino acid residues or residues involved in interaction with the chromophore has been successfully used to tune the spectra of other FPs [235, 236]. For example, a color transition from green to yellow was realized by introducing three mutations into zFP506, a close homolog of zFP538 sharing 87% identity [235]. Only 8 of the residues with sidechains directed into the interior of the protein are different between zFP506 and zFP538. The three mutations were the chromophore-forming residue Asn66 (mutated to Lys) and the two nearby residues Ala64 (mutated to Gly) and Asn69 (mutated to Asp). It should be noted that, like zFP538, zFP506 is an obligate tetramer. We chose mTFP1 as the recipient FP for the attempted transplantation of the zFP538-type chromophore [92]. The rationale for this choice was as follows: (1) mTFP1 is a well characterized monomeric FP which is bright, photostable and fast-folding [92]; (2) alignment of the crystal structure of zFP538 (PDB ID: 2OGR, chain A) and mTFP1 (PDB ID: 2HQB) reveals the overall backbone atom positions of these two FPs are exceptionally similar, with a 0.483 Å RMSD based on 158 equivalent C_α atoms (Figure 5.2); and (3) we have previously created a green-to-red photoconvertible FP using guided consensus design and mTFP1 as template [89], which suggests mTFP1 is amenable to substantial levels of residue substitution.

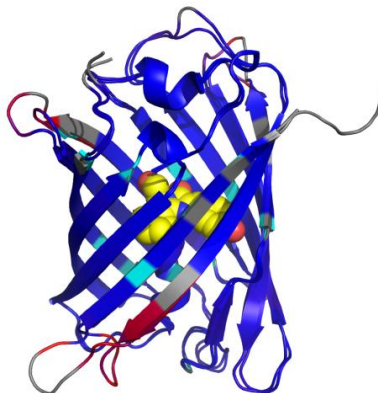


Figure 5.2 Alignment of mTFP1 (PDB ID: 2HQK) [92] and a single unit of tetrameric zFP538 (PDB ID: 2OGR, chain A) [74]. Blue specifies the minimum pairwise RMSD and red indicates the maximum while unaligned residues are coloured in gray. The overall RMSD based on 158 equivalent C_{α} atoms is 0.483 Å.

Previous studies have revealed that, in addition to Lys66, two other residues, Gly64 and Asp69, are critical for the formation of the 3rd ring in the chromophore zFP538 [235]. It was proposed that Gly64 provides spatial flexibility, while Asp69 plays a catalytic role in the transimination reaction [235, 237]. In addition, Ile44, together with nearby residues such as Leu46, Phe65, Leu204 and Leu219, forms a hydrophobic pocket that is thought to destabilize the positively charged N_{ξ} of Lys66 and thereby promote its cyclization [237]. After structural alignment and careful consideration of the residues with sidechains in close proximity to the chromophore, the following mutations were introduced into mTFP1 to create an 81 member library: Ile14Met, Leu16Tyr, Asn42Gln, Tyr58Glu, Thr62Thr/Ser/Ala, Thr63/Thr/Ser/Ala, **Ala64Gly**, **Ala66Lys**, **Asn69Asp**, Tyr120His, Ile199Ile/Leu/Val and Val213Ile/Leu/Val (the most critical mutations are highlighted in bold). Sequencing result of randomly picked variants confirmed that all of the targeted mutations were incorporated. In principle, variants in the library should have had identical chromophore-forming triad to zFP538 and, based on our models, the same chromophore microenvironment. However, none of the mutants in the library were either yellow or green fluorescent, implying the chromophore

was not properly formed. It is likely that some unidentified residues not at the primary proximity of the chromophore are also critical for successful formation of the distinct three-ring chromophore. The fact that zFP538 is the only naturally occurring FP with a KYG-derived yellow-emitting fluorophore also suggests that there is major barrier for such color transition. Consequently, we resorted to another strategy, which is to generate monomeric FP by starting from the zFP538 tetramer and systematically disrupting inter-subunit interactions.

5.2.2 Interface disruption of zFP538 and directed evolution

The most commonly used strategy to develop monomeric FPs from tetrameric precursors is to use mutagenesis to disrupt interactions between subunits in the tetramer. Specifically, mutations are introduced that switch formerly stabilizing interactions (i.e. close packing of hydrophobic residues, hydrogen bonds, or salt-bridges) into repulsive interactions (i.e. placing positively charged residues in close proximity). This approach, which relies on a combination of rational protein engineering and directed evolution, has been successfully used with a growing number of FPs [68, 87, 92, 96-99]. Given these successful examples, we proposed that a monomeric variant could be engineered from zFP538 by interface engineering and directed evolution.

Rational protein engineering is aided by the availability of high-resolution 3D structures of the protein of interest or a close homologue. Fortunately, the X-ray crystal structure of zFP538 has been determined [74, 237]. As shown in Figure 5.3, the zFP538 tetramer consists of two dimers packed at approximately right angles relative to each other, with each dimer formed by two head-to-tail packed monomers. Accordingly, there are two distinct types of interfaces in the tetramer. The head-to-tail interface (i.e. A-B and C-D interface in Figure 5.3) involves a large hydrophobic patch symmetrically centered on Ile106. Major interactions at this interface includes: van der Waals interactions between Leu98 (A) and Leu98 (B);

Val104 (A) and Val104 (B); Ile106 (A) and Ile106 (B); Met129 (A) and Val184 (B); and a hydrogen-bond interaction between Asn127 (A) and Asn127 (B) (Figure 5.4A) [74]. In contrast, the interactions at the right-angle interface (i.e. A-C and B-D interface in Figure 5.3) are more hydrophilic in nature and involve several ordered solvent molecules as observed from the crystal structure. Major interactions contributing to this interface include: hydrogen-bond interaction between Asp164 (A) and Tyr168 (C), Ser166 (A) and Asp164 (C), Gln180 (A) and Arg 178 (C); salt bridge interaction between Arg178 (A) and Asp164 (C); and van der Waals interaction between Cys149 (A) and Cys149 (C) (Figure 5.4B). Altogether, 8 hydrogen bonds, 4 salt bridge interactions, and a number of hydrophobic interactions (largely involving the last ~10 residues at the C-terminus of the protein) are observed at the A-C interface [74].

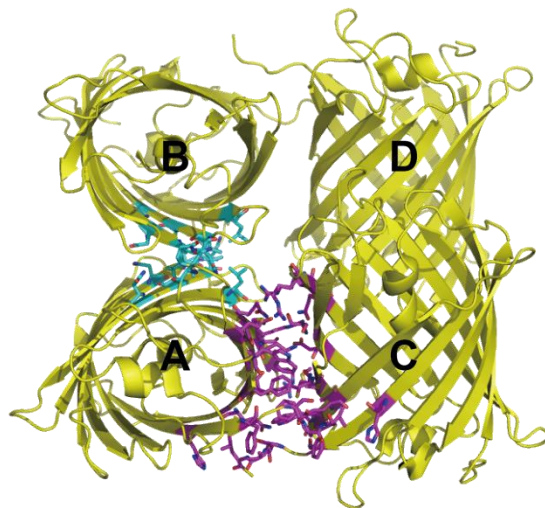


Figure 5.3 Cartoon representation of X-ray crystal structure of zFP538 (PDB ID: 2OGR) [74]. Two units (A and B, C and D) of zFP538 align anti-parallel to form a dimer. Most of the interactions at this head-to-tail interface are hydrophobic in nature. A tetramer is further formed by two dimers aligned at $\sim 90^\circ$. This right-angle interface is dominated by hydrophilic interactions. Residues contributing to the oligomer interaction are highlighted in cyan (for the head-to-tail A-B interface) or magenta (for the right-angle A-C interface).

To generate monomeric versions of zFP538, we first synthesized its encoding cDNA with codons optimized for mammalian expression. To

facilitate cloning, we replaced its first two residues (Met and Ala) with the “universal” FP N-terminal sequence (Met-Val-Ser-Lys-Gly-Glu-Glu), as well as appended the “universal” FP C-terminal sequence Met-Asp-Glu-Lys-Tyr-Lys to its C-terminus (Figure 5.7). These “universal” sequences correspond to the N- and C- terminal regions of avGFP and ensure that PCR primers designed to work with avGFP will also work for any other FP onto which these sequences have been grafted. Expressing this modified zFP538 in *E coli* resulted in yellow-fluorescing colonies. Screening of a gene library created by error-prone PCR resulted in the discovery of the brightness-enhancing mutation Met129Val. This mutation is also present in the commercially available zFP538 variant, known as ZsYellow (Clontech). Since the color of the purified protein of zFP538-M129V is very close to that of papaya fruit when viewed with naked eyes in white light, we named this variant as tetrameric Papaya version 0.01 (tPapaya0.01). Gel filtration chromatography (GFC) of purified tPapaya0.01 protein reveals that it exists as an equilibrating mixture of tetrameric (peak at 39 mL) and potentially dimeric species (peak at 58 mL) (Figure 5.5A).

In an effort to generate a monomeric zFP538, we initially attempted to disrupt the A-B interface by mutating Ile106 into Arg or Lys (Figure 5.4A). Fortunately, the resulting tPapaya0.01-I106R/V115E, named as dPapaya0.1, remained yellow fluorescent though it was substantially dimmer than tPapaya0.01 (Figure 5.5A). The GFC elution profile of the purified dPapaya0.1 confirmed that the protein was primarily a dimer (Figure 5.5B). To break the dimer into a monomer, we targeted Asp164 and Arg178 at the A-C interface with a subset of amino acids (Figure 5.4). The brightest variant in this small library, though still extremely dim (Figure 5.5A), was found to harbor Asp164Lys and Arg178His.

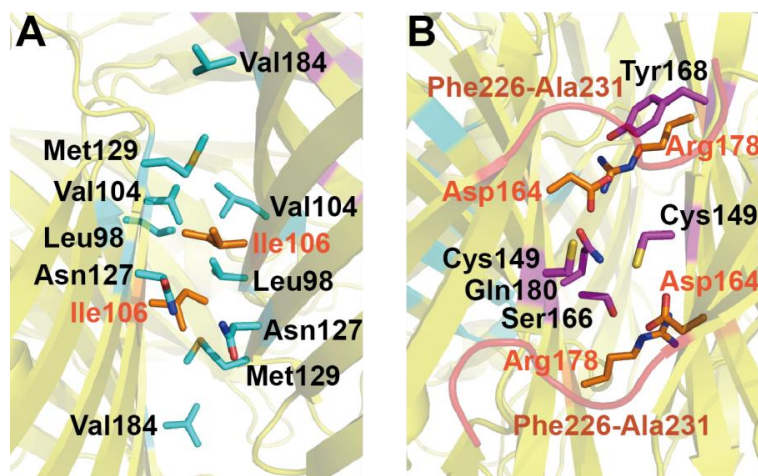


Figure 5.4 Zoom views of the inter-unit interfaces in zFP538 (PDB ID: 2OGR) [74]. (A) The hydrophobic A-B interface consisted of van der Waals interaction between Leu98, Val104 and Ile106, and hydrogen-bond interaction between Asn127. Disruption of the van der Waals interaction between Val106 (A) and Val106 (B) is sufficient to break the bright tetramer into a weakly yellow fluorescent dimer (Figure 5.5A and B). (B) The hydrophilic A-C interface which is majorly stabilized by hydrogen-bond interaction between Asp164 (A) and Tyr168 (C), Ser166 (A) and Asp164 (C), Gln180 (A) and Arg 178 (C), salt bridge interaction between Arg178 (A) and Asp164 (C), van der Waals interaction between Cys149 (A) and Cys149 (C). Double substitution of Asp164Lys and Arg178His generates a dim weak dimer (Figure 5.5A and B). The remaining dimer-forming tendency is significantly reduced by deleting the C-terminal residues interacting with each other (Phe226-Ala231).

To restore the brightness lost during monomerization, we performed 4 successive rounds of directed evolution by creating libraries by random mutagenesis and screening for brightness. This effort led to mPapaya0.27, which had 7 additional mutations and displayed reasonable fluorescence in *E. coli* after overnight incubation at 37°C (Figure 5.5A). Analysis of the purified protein by GFC revealed that the elution profile had two peaks: one at 72 mL and another at 65 mL with a shoulder at 58 mL (Figure 5.5B). Control experiments using size standards indicated that the peak at 72 mL likely corresponded to the monomeric species, while the peak at 65 mL was likely a dimer in dynamic equilibrium. The shoulder at 58 mL was likely a dynamic equilibrium of dimer and tetramer. This promising result suggested that mPapaya0.27 can exist as a monomer (presumably at low

enough concentrations) but can form dimers and tetramers at higher concentrations.

Since the last ten residues at the C-terminal of coral-derived FPs are known to promote oligomeric interaction (that is, they form a large part of the A-C interface), we next deleted the six residues from position 226 to 231 (Figure 5.4B, highlighted in red) and performed one more round of random mutagenesis and library screening. Truncation and modification at the C-terminus of eqFP611 has been previously reported to cause catastrophic effect on its folding [113]. Fortunately, we were able to find a variant, designated mPapaya0.3 that was as bright as the full-sequence mPapaya0.27 (Figure 5.5A). The GFC profile of the purified mPapaya0.3 revealed that the dimeric character was greatly reduced and it predominantly exists as a monomer (Figure 5.5B).

When the GFC elution profile was monitored at 280 nm, a wavelength at which all proteins absorb and is therefore a measure of total protein concentration, we found that a considerable portion of mPapaya0.3 eluted at 53 mL and 62 mL (Figure 5.5), suggesting the presence of tetrameric and dimeric oligomers, respectively, of mPapaya0.3. The difference in the elution profile when monitored at 280 nm and 525 nm indicates that the dimeric and tetrameric oligomers are composed of proteins that are inefficient at undergoing chromophore formation. Apparently, these proteins are sufficiently folded and soluble to be purified by affinity chromatography. We were concerned that the presence of these non-fluorescent oligomers could potentially interfere with the function and behavior of the protein of interest when mPapaya was used as fusion tag. Such interferences could contribute to increased cytotoxicity associated with introduction of the mPapaya gene. For this reason, we undertook further engineering in an effort to create a variant with further diminished oligomeric character.

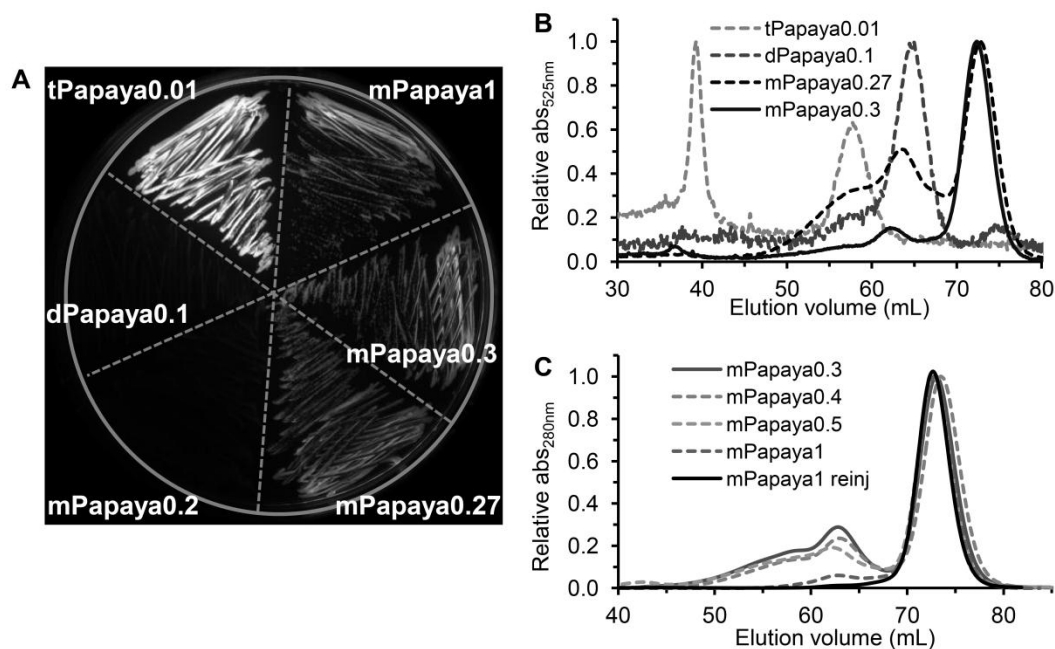


Figure 5.5 Fluorescent image and gel filtration chromatogram of Papaya variants. (A) Fluorescent image of *E. coli* expressing Papaya variants. *E. coli* was grown for 16 hours at 37°C before imaged. (B) GFC elution profiles of Papaya variants when the absorbance of the eluent is monitored at 525 nm. The monomeric FP mCherry eluted at 72 mL and the dimeric FP dTomato eluted at 67 mL were used as monomer and dimer standards, respectively (traces not shown). tPapaya exists as tetramer (eluted at 39 mL), with an additional component that is likely a dynamic equilibrium of tetramers and dimers (58 mL). dPapaya0.1 is a dimer (65 mL), while mPapaya0.27 presents as weak monomer (72 mL). mPapaya0.3 predominately exists as monomer. (B) GFC elution profiles of Papaya variants when the absorbance of the eluent is monitored at 280 nm. mPapaya0.3 has a significant peak at 65 mL with a shoulder at 53 mL. The dimer peak and tetramer shoulder are gradually reduced in subsequent variants. Reinjection of the monomer peak of mPapaya1 elutes exclusively as a monomer, suggesting that the small “dimer” component at 65 mL observed in the initial injection is an artifact.

Scrutinizing the protein-protein interfaces of the tetramer, as revealed in the X-ray crystal structure, led us to suspect that the weak oligomerization tendency could be due to the individual or cooperative result of the following interactions: Leu98(A)-Leu98(B), Val104(A)-Val104(B), Tyr127(A)-Tyr127(B), Val131(A)-Val(131B), Cys149(A)-Cys149(C) and Ser166(A)-Lys164(C). Accordingly, each of the above positions was mutated into Lys/Arg or Thr/Ser by site-directed mutagenesis and the

purified proteins were examined by GFC. Ultimately, we found the remaining oligomerization tendency was successfully abolished in mPapaya0.3 + Tyr127Arg/Cys149Thr/Ser166Lys, and so this variant was designated as mPapaya1. Of the three mutations, Ser166Lys is the main contributor to the decrease in oligomerization tendency based on the GFC elution profiles (Figure 5.5). Ser166(A) and Asp164(C) are engaged in a hydrogen bonding interaction in the A-C interface of wild-type zFP538 (Figure 5.4B). Introduction of a positively charged Lys residue at position 166 most likely introduces repulsive interactions with the side chain of Asp164Lys, which was introduced earlier in the engineering. A genealogy of all Papaya variants, as well as a map of the mutations is provided in Figure 5.6. Residues targeted for mutagenesis, the rationales and experimental results are summarized in Table D1 in the appendix. In addition, sequence alignment of mPapaya1 and some earlier variants is available in Figure 5.7.

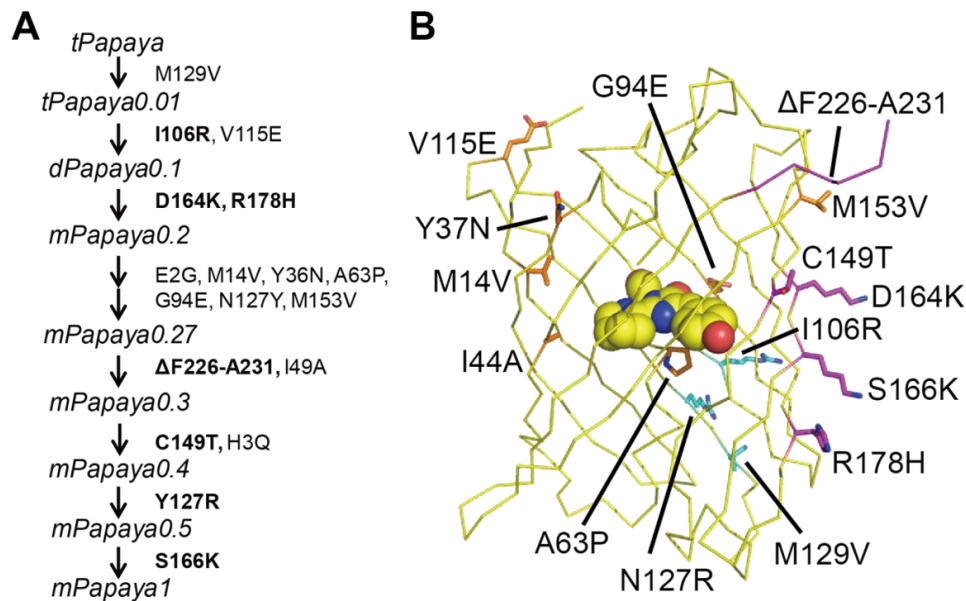


Figure 5.6 Overview of the directed evolution of mPapaya. (A) A genealogy of Papaya variants. (B) Locations of substitutions in mPapaya1 that were introduced during the directed evolution process. Mutations E2G and H3Q are not shown. The X-ray crystal structure of zFP538 (PDB ID: 2OGR, chain A) is used for representation [74].

	1	1a	1b	1c	1d	1e	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25	26	27	28	29	30	31	32	33	34	35	36	37	38	39	40	41	42	43	44	45	46	47	48	49	50	51	52	53	54	55
zFP538	M	-	-	-	-	A	H	S	K	H	G	L	K	E	E	M	T	M	K	Y	H	M	E	G	C	V	N	G	H	K	F	V	I	T	G	E	G	I	G	Y	P	F	K	G	K	Q	T	I	N	L	C	V	I	E	G	G	P	L	P	
tPapaya0.01	M	V	S	K	G	E	E	H	S	K	H	G	L	K	E	E	M	T	M	K	Y	H	M	E	G	C	V	N	G	H	K	F	V	I	T	G	E	G	I	G	Y	P	F	K	G	K	Q	T	I	N	L	C	V	I	E	G	G	P	L	P
dPapaya0.1	M	V	S	K	G	E	E	H	S	K	H	G	L	K	E	E	M	T	M	K	Y	H	M	E	G	C	V	N	G	H	K	F	V	I	T	G	E	G	I	G	Y	P	F	K	G	K	Q	T	I	N	L	C	V	I	E	G	G	P	L	P
mPapaya0.3	M	V	S	K	G	E	G	H	S	K	H	G	L	K	E	E	M	T	V	K	Y	H	M	E	G	C	V	N	G	H	K	F	V	I	T	G	E	G	I	G	Y	P	F	K	G	K	Q	T	A	N	L	C	V	I	E	G	G	P	L	P
mPapaya1	M	V	S	K	G	E	G	Q	S	K	H	G	L	K	E	E	M	T	V	K	Y	H	M	E	G	C	V	N	G	H	K	F	V	I	T	G	E	G	I	G	N	P	F	K	G	K	Q	T	A	N	L	C	V	I	E	G	G	P	L	P
zFP538	56	57	58	59	60	61	62	63	64	65	66	67	68	69	70	71	72	73	74	75	76	77	78	79	80	81	82	83	84	85	86	87	88	89	90	91	92	93	94	95	96	97	98	99	100	101	102	103	104	105	106	107	108	109	110	111	112	113	114	115
tPapaya0.01	F	S	E	D	I	L	S	A	G	F	K	Y	G	D	R	I	F	T	E	Y	P	Q	D	I	V	D	Y	F	K	N	S	C	P	A	G	Y	T	W	G	R	S	F	L	F	E	D	G	A	V	C	I	C	N	V	D	I	T	V	S	V
dPapaya0.1	F	S	E	D	I	L	S	A	G	F	K	Y	G	D	R	I	F	T	E	Y	P	Q	D	I	V	D	Y	F	K	N	S	C	P	A	G	Y	T	W	G	R	S	F	L	F	E	D	G	A	V	C	R	C	N	V	D	I	T	V	S	E
mPapaya0.3	F	S	E	D	I	L	S	A	G	F	K	Y	G	D	R	I	F	T	E	Y	P	Q	D	I	V	D	Y	F	K	N	S	C	P	A	G	Y	T	W	E	R	S	F	L	F	E	D	G	A	V	C	R	C	N	V	D	I	T	V	S	E
mPapaya1	F	S	E	D	I	L	S	A	G	F	K	Y	G	D	R	I	F	T	E	Y	P	Q	D	I	V	D	Y	F	K	N	S	C	P	A	G	Y	T	W	E	R	S	F	L	F	E	D	G	A	V	C	R	C	N	V	D	I	T	V	S	E
zFP538	116	117	118	119	120	121	122	123	124	125	126	127	128	129	130	131	132	133	134	135	136	137	138	139	140	141	142	143	144	145	146	147	148	149	150	151	152	153	154	155	156	157	158	159	160	161	162	163	164	165	166	167	168	169	170	171	172	173	174	175
tPapaya0.01	K	E	N	C	I	Y	H	K	S	I	F	N	G	V	N	F	P	A	D	G	P	V	M	K	K	M	T	T	N	W	E	A	S	C	E	K	I	M	P	V	P	K	Q	G	I	L	K	G	D	V	S	M	Y	L	L	L	K	D	G	G
dPapaya0.1	K	E	N	C	I	Y	H	K	S	I	F	N	G	V	N	F	P	A	D	G	P	V	M	K	K	M	T	T	N	W	E	A	S	C	E	K	I	M	P	V	P	K	Q	G	I	L	K	G	D	V	S	M	Y	L	L	L	K	D	G	G
mPapaya0.3	K	E	N	C	I	Y	H	K	S	I	F	Y	G	V	N	F	P	A	D	G	P	V	M	K	K	M	T	T	N	W	E	A	S	C	E	K	I	V	P	V	P	K	Q	G	I	L	K	G	K	V	S	M	Y	L	L	L	K	D	G	G
mPapaya1	K	E	N	C	I	Y	H	K	S	I	F	R	G	V	N	F	P	A	D	G	P	V	M	K	K	M	T	T	N	W	E	A	S	T	E	K	I	V	P	V	P	K	Q	G	I	L	K	G	K	V	K	M	Y	L	L	L	K	D	G	G
zFP538	176	177	178	179	180	181	182	183	184	185	186	187	188	189	190	191	192	193	194	195	196	197	198	199	200	201	202	203	204	205	206	207	208	209	210	211	212	213	214	215	216	217	218	219	220	221	222	223	224	225	226	227	228	229	230	231	232	233	234	235
tPapaya0.01	R	Y	R	C	Q	F	D	T	V	Y	K	A	K	S	V	P	S	K	M	P	E	W	H	F	I	Q	H	K	L	L	R	E	D	R	S	D	A	K	N	Q	K	W	Q	L	T	E	H	A	I	A	F	P	S	A	L	A	-	-	-	
dPapaya0.1	R	Y	R	C	Q	F	D	T	V	Y	K	A	K	S	V	P	S	K	M	P	E	W	H	F	I	Q	H	K	L	L	R	E	D	R	S	D	A	K	N	Q	K	W	Q	L	T	E	H	A	I	A	F	P	S	A	L	A	-	-	-	
mPapaya0.3	R	Y	H	C	Q	F	D	T	V	Y	K	A	K	S	V	P	S	K	M	P	E	W	H	F	I	Q	H	K	L	L	R	E	D	R	S	D	A	K	N	Q	K	W	Q	L	T	E	H	A	I	A	-	-	-	-	-	-	G	M	D	E
mPapaya1	R	Y	H	C	Q	F	D	T	V	Y	K	A	K	S	V	P	S	K	M	P	E	W	H	F	I	Q	H	K	L	L	R	E	D	R	S	D	A	K	N	Q	K	W	Q	L	T	E	H	A	I	A	-	-	-	-	-	-	G	M	D	E
zFP538	236	237	238																																																									
tPapaya0.01	-	-	-																																																									
dPapaya0.1	L	Y	K																																																									
mPapaya0.3	L	Y	K																																																									
mPapaya1	L	Y	K																																																									

Figure 5.7 Sequence alignment of zFP538 and Papaya variants. The chromophore forming triad Lys-Tyr-Gly is highlighted in orange color and boxed. Positions targeted for semi-saturated mutagenesis are highlighted in box. Mutations discovered in this work is highlighted in grey shape.

5.2.3 Validation of mPapaya1 as a monomer under physiological conditions

Although the elution profile of mPapaya1 shows that it exists as a monomer *in vitro*, we were concerned it might still oligomerize under certain physiological condition as has been observed in some other monomeric FPs that were claimed to be monomeric [114, 238]. Some monomeric FPs, when fused to membrane-associated or naturally oligomeric proteins, form oligomers or higher order aggregates due to the elevated effective concentration at a confined geometry. Accordingly, we generated several membrane targeted fusions of mPapaya1, namely CytERM-mPapaya1 and Orai1-mPapaya1, expressed them in mammalian cells, and compared their localization pattern with those of well-known monomeric FPs. These two fusion partners were chosen because they are membrane proteins known to represent particularly “challenging” targets for FP fusions [101, 238]. The CytERM is an endoplasmic reticulum (ER) signal-peptide which anchors its downstream protein to the cytoplasmic side of ER. It was previously reported that the low affinity interactions between the cytoplasmic domains of ER membrane resident proteins can restructure the tubular ER structure into stacked membrane arrays (i.e. organized smooth ER (OSER)) structure [239]. For instance, bright whorl-like OSER structures were frequently observed in cells expressing CytERM-EGFP, but not for CytERM-mEGFP (Figure 5.8A) [239]. In addition, the CytERM fusion of TagRed, which was reported to be exist as a monomer at high concentration (0.4 mM) based on gel filtration data [114], was found to exhibit more whorl-like structures than CytERM-EGFP [238].

In our own control experiments, we observed approximate 1/3 of the cells expressing CytERM-mCherry have extensive OSER whorls (Figure 5.8C). In contrast, the majority of the cells expressing CytERM-mPapaya1 show an evenly distributed reticular ER structure (Figure 5.8B),

comparable to what is observed in cells expressing CytERM-mEGFP (Figure 5.8A). Orai1, a plasma membrane protein that exists as homodimer at rest state and forms tetramer when activated [240], stands as another “challenging” target for FP fusions. For the Orai1 fusions, mPapaya1 qualitatively performs better than YFP, which is a known monomeric FP, and displayed a distinctive pattern of plasma membrane localization, with relatively few fluorescent aggregating structures apparent inside the cytoplasm. Moreover, when fused to β -actin, mPapaya1 correctly reports the fine filament structures formed by monomeric actin subunits in a highly ordered arrangement. These results suggest that mPapaya1 remains monomeric under physiological conditions.

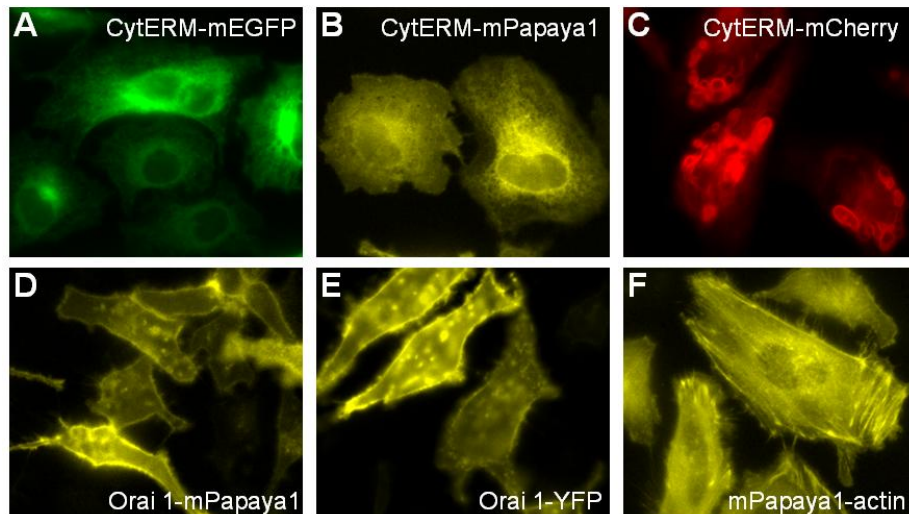


Figure 5.8 Representative widefield fluorescence images of HeLa cells expressing FP-fusions. (A-C) Majority cells expressing CytERM-mEGFP (A) and CytERM-mPapaya (B) show tubular ER structure. In contrast, cells expressing CytERM-mCherry (C) displays extensive OSER whorls. (D-E) Both Orai1-mPapaya1 (D) and Orai1-YFP (E) expressing cells display distinct plasma membrane localization. However, aggregating bright spots is more frequently observed in cells expressing Orai1-YFP. (F) Cells expressing mPapaya1- β -actin, the fine filament structure of actin is well preserved. (C) is a courtesy of Yi Shen.

5.2.4 *In vitro* characterization of mPapaya1

Having confirmed that mPapaya1 is a true monomer, we continued to perform *in vitro* characterization. Similar to zFP538, mPapaya1 has an excitation maximum at 529 nm and a slightly red-shifted emission maximum at 541 nm (Figure 5.9A and Table 5.1). There is a non-excitable absorbance peak, with a maximum at 415 nm, which is attributable to the protonated state of the chromophore. A fluorescence vs. pH titration reveals that mPapaya1 has a relatively high pK_a of 6.8 (Figure 5.9B). Studying the absorbance spectra of earlier mPapaya variants reveals that the absorbance of the protonated state was negligible until Cys149Thr was introduced at the stage of mPapaya0.4 (Figure 5.6A). Indeed, mPapaya0.3 has a lower pK_a of 6.4. In an attempt to lower the pK_a of mPapaya1, we replaced Thr149 with Ser/Val/Ala. Unfortunately, the pK_a of both T149S and T149A remained 6.8, while T149V exhibited much dimmer fluorescence. Despite its higher pK_a , mPapaya1 has 177% of the brightness than ZsYellow1 at physiological pH (7.4). The increased brightness is largely attributed to an increased extinction coefficient.

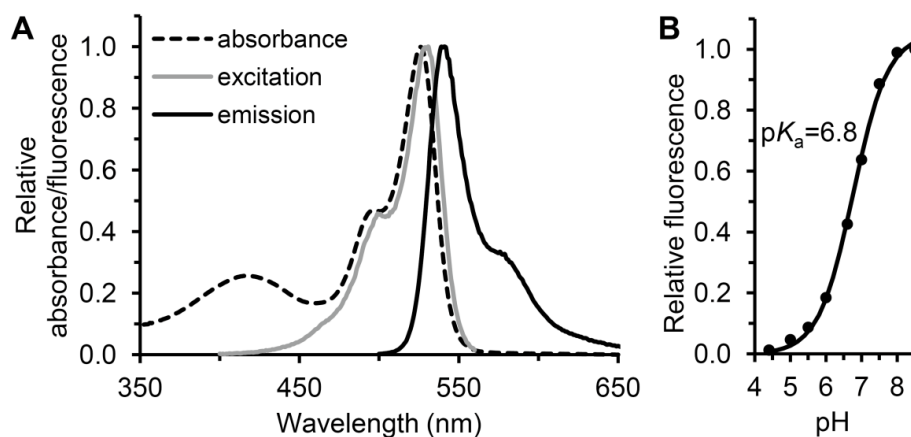


Figure 5.9 Spectral profile and pH titration of mPapaya1. (A) Absorbance, excitation and emission spectra of mPapaya1. (B) Fraction of the deprotonated (fluorescent) chromophore as a function of pH. Data is well fit to Hill equation with a Hill coefficient of 0.89.

Table 5.1 Properties of mPapaya1 and ZsYellow1

Protein	λ_{exc} (nm)	λ_{em} (nm)	ϵ^{a}	Φ	Brightness ^b	pK _a	Oligomeric structure
mPapaya1	529	541	35	0.65	23	6.8	monomer
ZsYellow1 ^c	529	539	20	0.65	13	ND	tetramer

^a Unit of $\text{mM}^{-1} \text{cm}^{-1}$. ^b Product of ϵ and Φ in $\text{mM}^{-1} \text{cm}^{-1}$. ^c Data from Clontech website (<http://www.clontech.com/>).

5.3 Conclusions

In an attempt to create a true yellow FP with a cyclic-imine substituted p-HBI chromophore and emission maximum around 538 nm, we initially tried to transplant the chromophore and its environment of zFP538 into mTFP1, a well-characterized monomeric FP. This approach was ultimately unsuccessful, as the extensive amount of sequence modification rendered the mTFP1 variant non-fluorescent. In the second approach, we use a combination of semi-rational protein engineering and directed evolution to turn zFP538 into a monomer. Semi-rationally introducing charged residues mediating repulsive interactions at the inter-subunit interfaces successfully broke down the interfaces and ultimately led to monomeric variants. Subsequently, the brightness lost during monomerization was rescued by directed evolution. The final product, designated as mPapaya1, was demonstrated to be a monomer *in vitro* and *in vivo* and have an intrinsic brightness that is 1.7-fold of the commercially available ZsYellow. The emission maximum of mPapaya1 (541 nm) fills the current spectra gap between monomeric yellow FPs and monomeric orange FPs, presenting new opportunities for multi-color imaging and FRET pair construction, though such applications have not yet been demonstrated. Unfortunately, we did notice that mPapaya has relatively fast photobleaching kinetics, which could be problematic for certain time-lapse imaging experiments. We expect more photostable variants can be developed by continuing directed evolution for improved photostability.

5.4 Materials and methods

5.4.1 General methods and materials

See Section 2.4.1. The mammalian expression plasmid encoding CytERM-mEGFP is a courtesy from Professor Erik Snapp at Albert Einstein College of Medicine. The plasmid encoding Orai-YFP is a courtesy from Professor Patrick Hogan at Northwestern University (acquired through Addgene, plasmid number 19756).

5.4.2 Plasmid library creation and screening

For plasmid library creation, general molecular biology experiments were performed similarly as described in Section 2.4.2 unless else specified. Saturation or semi-saturation mutagenesis at specific residues was performed using overlap extension PCR or QuickChange Site-Directed Mutagenesis Kit (Agilent Technologies). *E. coli* libraries was screened for brightness using the following filters: 490-510 nm for excitation and 520-550 nm for emission.

5.4.3 Protein purification and characterization

Protein purification, pH titration and gel filtration chromatography were performed similarly as described in Section 2.4.4 unless else specified. The extinction coefficient was calculated by divided the absorbance by the protein concentration, which is in turn determined by BCA protein assay kit (Pierce). The yellow FP mCitrine ($\Phi = 0.76$) was used as reference for quantum yield determination [105, 229].

5.4.4 Live cell imaging

Construction of mammalian expression plasmid. To construct Orai1-mPapaya1 and CytERM-mPapaya1, the gene of mPapaya1 was amplified with a 5' primer with a *AgeI* site and a 3' primer with a *NotI* site. The

purified PCR products were digested and ligated into a similarly digested pOrai1-YFP or pCytERM-mEGFP (see Section 5.4.1 for source of plasmid). To create the pmPapaya1-actin constructs, we amplified the gene encoding mPapaya1 with a 5' primer with an *NheI* site and a 3' primer with an *XhoI* site. The purified PCR products were then digested and ligated into a similarly digested pEGPF-actin (Clontech).

Cell maintenance and transient transfection. See Section 3.3.3. The same Nikon Eclipse Ti microscope was used. Fluorescence of mPapaya1 was observed using the conventional FITC filter set (i.e. 483-493 nm (excitation), 500-550 nm (emission), and 488 nm (dichroic)).

Chapter 6 Conclusions and future directions

6.1 Summary of the thesis

The oligomeric nature of Anthozoa FPs is detrimental to most applications of FPs [87]. Accordingly, generation of monomeric variants is a *prerequisite* step to create FPs practically useful for fluorescent labeling and biosensor development. The focus of this thesis was to develop new monomeric FPs with unique properties and expand their applications as biosensing tools.

In the past several years, photoconvertible FPs (pcFPs) have attracted significant attention in the field of FP engineering due to their “highlightable” nature that allows researchers to study protein kinetics and acquire sub-diffraction-limit images [128]. Starting with mClavGR1, a green-to-red pcFP created by consensus design using the monomeric mTFP1 as template, I performed directed evolution for improved photoconversion efficiency using a custom-built illumination source and imaging system. This effort ultimately led to the production of mClavGR2. Subsequent optimization for monomeric behavior and photoconversion efficiency led to the discovery of mMaple. Characterization results demonstrated that both mClavGR2 and mMaple have favorable spectral properties and retain the monomeric structure of the progenitor FP, mTFP1. Improved photoconversion efficiency achieved in the mClavGR series is the result of shifting the phenol/phenolate equilibrium in the pre-photoconversion green species towards the phenol state. I have demonstrated that mClavGR2 is useful for protein trafficking studies, as well as for protein kinetics studies. Furthermore, in close collaboration with Ann McEvoy and other researchers, we have demonstrated that mMaple can be used as a multi-modal probe for both (f-)PALM/STORM and SIM, two different super-resolution microscopy techniques. The maturation, folding, photoswitching and photostability of mMaple and mEos2 (a widely used pcFP), were characterized and compared. Based on the results, we concluded that the improved folding and photostability led to the superior

performance of mMaple in certain super-resolution imaging experiments. The detailed characterization of mMaple will also provides insight and serve as useful benchmarks for researchers who are interested in using mMaple or other pcFPs to decipher biological phenomena.

Another important implementation of FPs lies in their application as genetically encoded biosensors [241]. We envisioned that combining the “highlightable” asset of photomodulable FPs and the Ca^{2+} -sensing capability of genetically encoded Ca^{2+} indicators into one single construct would produce a highlightable Ca^{2+} indicator. Such a hybrid tool could allow researchers to spectrally mark a sub-population of cells in a multicellular system and subsequently monitor their Ca^{2+} dynamics over space and time. I explored four different protein design strategies to create highlightable Ca^{2+} indicators. The strategy that ultimately proved successful involved the creation of a circularly permuted version of mMaple and its incorporation into a G-CaMP-type single FP-based Ca^{2+} indicator. Subsequent optimization by extensive semi-rational protein engineering and directed evolution led to the identification of two promising variants, designated as GR-GECO1.1 and GR-GECO1.2, that exhibit excellent photoconversion properties and have an up to 4.6-fold increase in red fluorescence intensity upon binding Ca^{2+} . The spectral properties, mechanism, and kinetics of these novel Ca^{2+} indicators were studied. I have shown that single cells (cervical cancer cells and hippocampal neuron cells) expressing the GR-GECOs can be highlighted using near UV illumination and both the green and the red fluorescence emission channels are reflective of the concentration of intracellular Ca^{2+} . We expect that GR-GECOs will be particularly useful in applications such as dissecting Ca^{2+} signaling processes in multicellular system or during cell development.

In Chapter 5, I described the successful engineering of another monomeric FP, named as mPapaya1. Acquired by step-by-step breaking

down of the inter-subunit interfaces of the Anthozoan FP zFP538, mPapaya1 represents the first monomeric FP with a cyclicimine substituted p-HBI chromophore and gives true yellow emission. The oligomeric status of mPapaya1 is characterized and its monomeric status confirmed both in vitro and under physiological conditions. Spectral properties of mPapaya1 were also studied and discussed.

6.2 Future works

6.2.1 The photoconvertible FPs mClavGR2 and mMaple

For future work, it would be an asset if the photoswitching properties of mClavGR2 and mMaple can be both removed, and enhanced, in future variants. Removing the reversible photoswitching tendency would reduce the complexity of their photophysics and increase the accuracy of single-molecule counting in the (f-)PALM/STORM application. On the other hand, enhancing this property would lead to a photochromic variant that can be irreversibly photoconverted from green species to red species and reversibly photoswitched between dark and fluorescent states at each of the two species. Indeed, assisted by the structural insight provided by crystal structures, two such photochromic FPs, namely mIris and NijiFP, were rationally engineered from mEos and Dendra2 respectively [242, 243]. As discussed in Chapter 2, all of the pcFPs display a large extent of sequence and functionality similarity. Accordingly, we hypothesize that the photochromic tendency of mMaple could be enhanced by introducing mutations similar to those found in mIris and NijiFP. Interestingly, in our initial test, the F177S substitution (i.e. the only substitution in NijiFP and the key mutation in mIrisFP) into mMaple did not provide any improvement on its photoswitching property. Further engineering would be facilitated by the availability of a crystal structure.

6.2.2 The highlightable Ca²⁺ indicators GR-GECOs

Additional investigation on the Ca²⁺-binding kinetics of GR-GECO1.2 is required as the stopped-flow spectroscopic data suggested that it is Ca²⁺ concentration independent. This result is difficult to rationalize, therefore understanding the responding mechanism may help avoid the misinterpretation of results when the indicators are used to investigate biological phenomena. Even though extensive engineering have been done to the GR-GECOs, their dynamic range remains moderate compared with other members in the GECO family [168]. This suggests that there might exist some fundamental barrier similar to the trade-off between the brightness of the green state and the photoconversion efficiency in mClavGR2 and mMaple. We noticed that GR-GECOs retain the high pKa characteristic of mMaple (e.g. 8.5 for Ca²⁺-bound green state and 7.8 for Ca²⁺-bound red state of GR-GECO1.1) which is likely the result of the two-step screening process. That is, libraries were screened for photoconversion efficiency prior to screening for Ca²⁺ response. Such a high pKa characteristic not only renders the signal interpretation more complicated (i.e. pH fluctuation need to be monitored), but also limits the dynamic range at physiological pH. Of the monomeric green-to-red pcFPs known to date, mEos2 has the lowest pKa values (5.6 for the green state and 6.4 for the red state) and it undergoes photoconversion even though its green chromophore exists predominantly in the anionic state. Replacing cpMaple in GR-GECO with cp variants of mEos2 or its improved mutant mEos3.1 [101] might led to GR-GECO variants with improved dynamic range. However, risk remains as circularly permuted variants of mEos2 might fail to fold properly since the folding mEos2 is known to be sub-optimal. Also, continuing pursuing directed evolution of paR-GECO and psO-GECO based on the two-step screening strategy might ultimately yield useful constructs. Again, availability of crystal structures would

provide more insight into the acting mechanism and thus offers new clues for engineering.

6.2.3 The monomeric FP with a true yellow emission mPapaya1

Experimental validation of mPapaya1's potential application as a FRET pair partner--either being an acceptor for green- and cyan-hue FPs such as EGFP and mTFP1, or being a donor for red FPs such as mRuby2 [151] is an obvious future direction for this work. In addition, we have noticed that mPapaya1 undergoes rapid photobleaching and this is undesirable for most live cell imaging applications. Initial studies revealed that blue/green excitation light drives the fluorescent state into a dark state that has absorbance maximum at around 400 nm and can return to the fluorescent state in a time scale of minutes. This recoverable fast photobleaching is reminiscent of the photobleaching behavior of mTFP0.7, whose fast photobleaching is indeed a photo-induced *trans-cis* isomerization of the chromophore [72]. Since we have made extensive modification on the β -barrel to break the oligomeric interfaces, it is likely some minor modification on the backbone of the barrel strands alternate the micro-environment of the chromophore. Introduction of mutations that can restraint the flexibility of the chromophore should be able to improve the photostability of mPapaya1.

6.3 Closing remark

In the research work described in this thesis, we have successfully created new monomeric FPs with distinct spectral properties and used directed evolution to evolve variants with improved performance on properties such as photoconversion efficiency, brightness and photostability. We also noticed that there are certain limitations in the new FPs and the FP-derived biosensors. To finish, I would like to reiterate that while the FP technology has become an established and trusted tool, none

of the FPs can be claimed to be flawless. Though extensive engineering effort has been invested and we do have a wide selection of engineered FPs with superior performance at desired properties, limitations can still be found in these improved variants. Indeed, as FPs are getting widely used, and numerous data been collected to interpret biological problems, some shortfalls of FPs that we were previously not aware of have come to light. Questions or drawbacks about FP that have emerged during their applications include the complex photophysics of excited states [244-248], non-linear relationship between photobleaching kinetics and illumination intensity [84], context-dependent photobleaching behavior [81, 249], incomplete maturation [250], cell toxicity [251], and others. These pitfalls can lead to experimental artefacts and mis-interpretation of the biological phenomena under investigation. In addition, many biosensors are reported to perform less optimal in transfected tissues or transgenic animals than in *in vitro* cell culture [252]. Overcoming these emerging limitations of FPs demands continuing effort from protein engineers and a new generation of FPs that come another stop closer to being ideal.

Bibliography

- (1) Tsien, R. The green fluorescent protein. *Annu. Rev. Biochem.* **1998**, *67*, 509-544.
- (2) Shaner, N. C.; Patterson, G. H.; Davidson, M. W. Advances in fluorescent protein technology. *J. Cell. Sci.* **2007**, *120*, 4247-4260.
- (3) Davidson, M. W.; Campbell, R. E. Engineered fluorescent proteins: innovations and applications. *Nat. Methods* **2009**, *6*, 713-717.
- (4) Shaner, N. C.; Steinbach, P. A.; Tsien, R. Y. A guide to choosing fluorescent proteins. *Nat. Methods* **2005**, *2*, 905-909.
- (5) Day, R. N.; Davidson, M. W. The fluorescent protein palette: tools for cellular imaging. *Chem. Soc. Rev.* **2009**, *38*, 2887-2921.
- (6) Campbell, R. E. Fluorescent-protein-based biosensors: modulation of energy transfer as a design principle. *Anal. Chem.* **2009**, *81*, 5972-5979.
- (7) Boas, F. E.; Harbury, P. B. Potential energy functions for protein design. *Curr. Opin. Struct. Biol.* **2007**, *17*, 199-204.
- (8) Sternberg, M. J.; Thornton, J. M. Prediction of protein structure from amino acid sequence. *Nature* **1978**, *271*, 15-20.
- (9) Moulton, J.; Pedersen, J. T.; Judson, R.; Fidelis, K. A large-scale experiment to assess protein structure prediction methods. *Proteins* **1995**, *23*, ii-v.
- (10) Cordes, M.; Davidson, A.; Sauer, R. Sequence space, folding and protein design. *Curr. Opin. Struct. Biol.* **1996**, *6*, 3-10.
- (11) Dahiyat, B. I.; Mayo, S. L. De novo protein design: fully automated sequence selection. *Science* **1997**, *278*, 82-87.
- (12) Kuhlman, B.; Dantas, G.; Ireton, G. C.; Varani, G.; Stoddard, B. L.; Baker, D. Design of a novel globular protein fold with atomic-level accuracy. *Science* **2003**, *302*, 1364-1368.
- (13) Pantazes, R. J.; Grisewood, M. J.; Maranas, C. D. Recent advances in computational protein design. *Curr. Opin. Struct. Biol.* **2011**, *21*, 467-472.
- (14) Reina, J.; Lacroix, E.; Hobson, S. D.; Fernandez-Ballester, G.; Rybin, V.; Schwab, M. S.; Serrano, L.; Gonzalez, C. Computer-aided design of a PDZ domain to recognize new target sequences. *Nat. Struct. Biol.* **2002**, *9*, 621-627.

- (15) Chevalier, B. S.; Kortemme, T.; Chadsey, M. S.; Baker, D.; Monnat, R. J.; Stoddard, B. L. Design, activity, and structure of a highly specific artificial endonuclease. *Mol. Cell* **2002**, *10*, 895-905.
- (16) Jiang, L.; Althoff, E. A.; Clemente, F. R.; Doyle, L.; Rothlisberger, D.; Zanghellini, A.; Gallaher, J. L.; Betker, J. L.; Tanaka, F.; Barbas, C. F., 3rd; Hilvert, D.; Houk, K. N.; Stoddard, B. L.; Baker, D. De novo computational design of retro-aldol enzymes. *Science* **2008**, *319*, 1387-1391.
- (17) Siegel, J. B.; Zanghellini, A.; Lovick, H. M.; Kiss, G.; Lambert, A. R.; St Clair, J. L.; Gallaher, J. L.; Hilvert, D.; Gelb, M. H.; Stoddard, B. L.; Houk, K. N.; Michael, F. E.; Baker, D. Computational design of an enzyme catalyst for a stereoselective bimolecular Diels-Alder reaction. *Science* **2010**, *329*, 309-313.
- (18) Rothlisberger, D.; Khersonsky, O.; Wollacott, A. M.; Jiang, L.; DeChancie, J.; Betker, J.; Gallaher, J. L.; Althoff, E. A.; Zanghellini, A.; Dym, O.; Albeck, S.; Houk, K. N.; Tawfik, D. S.; Baker, D. Kemp elimination catalysts by computational enzyme design. *Nature* **2008**, *453*, 190-195.
- (19) Baker, D. An exciting but challenging road ahead for computational enzyme design. *Protein Sci.* **2010**, *19*, 1817-1819.
- (20) Tobin, M. B.; Gustafsson, C.; Huisman, G. W. Directed evolution: the 'rational' basis for 'irrational' design. *Curr. Opin. Struct. Biol.* **2000**, *10*, 421-427.
- (21) Schmidt-Dannert, C. Directed evolution of single proteins, metabolic pathways, and viruses. *Biochemistry* **2001**, *40*, 13125-13136.
- (22) Chen, K.; Arnold, F. H. Tuning the activity of an enzyme for unusual environments: sequential random mutagenesis of subtilisin E for catalysis in dimethylformamide. *Proc. Natl. Acad. Sci. U. S. A.* **1993**, *90*, 5618-5622.
- (23) Romero, P. A.; Arnold, F. H. Exploring protein fitness landscapes by directed evolution. *Nat. Rev. Mol. Cell Biol.* **2009**, *10*, 866-876.
- (24) Turner, N. J. Directed evolution drives the next generation of biocatalysts. *Nat. Chem. Biol.* **2009**, *5*, 567-573.
- (25) Chica, R. A.; Doucet, N.; Pelletier, J. N. Semi-rational approaches to engineering enzyme activity: combining the benefits of directed evolution and rational design. *Curr. Opin. Biotechnol.* **2005**, *16*, 378-384.
- (26) Fox, R. J.; Huisman, G. W. Enzyme optimization: moving from blind evolution to statistical exploration of sequence-function space. *Trends Biotechnol.* **2008**, *26*, 132-138.

- (27) Chica, R. A.; Moore, M. M.; Allen, B. D.; Mayo, S. L. Generation of longer emission wavelength red fluorescent proteins using computationally designed libraries. *Proc. Natl. Acad. Sci. U. S. A.* **2010**, *107*, 20257-20262.
- (28) Fox, R. J.; Davis, S. C.; Mundorff, E. C.; Newman, L. M.; Gavrilovic, V.; Ma, S. K.; Chung, L. M.; Ching, C.; Tam, S.; Muley, S.; Grate, J.; Gruber, J.; Whitman, J. C.; Sheldon, R. A.; Huisman, G. W. Improving catalytic function by ProSAR-driven enzyme evolution. *Nat. Biotechnol.* **2007**, *25*, 338-344.
- (29) Hutchison, C. A., 3rd; Phillips, S.; Edgell, M. H.; Gillam, S.; Jahnke, P.; Smith, M. Mutagenesis at a specific position in a DNA sequence. *J. Biol. Chem.* **1978**, *253*, 6551-6560.
- (30) Kunkel, T. A.; Loeb, L. A. On the fidelity of DNA replication. Effect of divalent metal ion activators and deoxyrionucleoside triphosphate pools on in vitro mutagenesis. *J. Biol. Chem.* **1979**, *254*, 5718-5725.
- (31) Cirino, P. C.; Mayer, K. M.; Umeno, D. Generating mutant libraries using error-prone PCR. *Methods Mol. Biol.* **2003**, *231*, 3-9.
- (32) Zaccolo, M.; Gherardi, E. The effect of high-frequency random mutagenesis on in vitro protein evolution: a study on TEM-1 beta-lactamase. *J. Mol. Biol.* **1999**, *285*, 775-783.
- (33) Stemmer, W. P. Rapid evolution of a protein in vitro by DNA shuffling. *Nature* **1994**, *370*, 389-391.
- (34) Stemmer, W. P. Dna Shuffling by Random Fragmentation and Reassembly - In-Vitro Recombination for Molecular Evolution. *Proc. Natl. Acad. Sci. U. S. A.* **1994**, *91*, 10747-10751.
- (35) Cramer, A.; Raillard, S. A.; Bermudez, E.; Stemmer, W. P. DNA shuffling of a family of genes from diverse species accelerates directed evolution. *Nature* **1998**, *391*, 288-291.
- (36) Shao, Z.; Zhao, H.; Giver, L.; Arnold, F. H. Random-priming in vitro recombination: an effective tool for directed evolution. *Nucleic Acids Res.* **1998**, *26*, 681-683.
- (37) Zhao, H.; Giver, L.; Shao, Z.; Affholter, J. A.; Arnold, F. H. Molecular evolution by staggered extension process (StEP) in vitro recombination. *Nat. Biotechnol.* **1998**, *16*, 258-261.
- (38) Zhao, H.; Arnold, F. H. Combinatorial protein design: strategies for screening protein libraries. *Curr. Opin. Struct. Biol.* **1997**, *7*, 480-485.
- (39) Lin, H.; Cornish, V. W. Screening and selection methods for large-scale analysis of protein function. *Angew. Chem. Int. Ed.* **2002**, *41*, 4402-4425.

- (40) Shimomura, O.; Johnson, F. H.; Saiga, Y. Extraction, purification and properties of aequorin, a bioluminescent protein from the luminous hydromedusan, *Aequorea*. *J. Cell. Comp. Physiol.* **1962**, *59*, 223-239.
- (41) Chalfie, M.; Tu, Y.; Euskirchen, G.; Ward, W. W.; Prasher, D. C. Green fluorescent protein as a marker for gene expression. *Science* **1994**, *263*, 802-805.
- (42) Matz, M. V.; Fradkov, A. F.; Labas, Y. A.; Savitsky, A. P.; Zaraisky, A. G.; Markelov, M. L.; Lukyanov, S. A. Fluorescent proteins from nonbioluminescent Anthozoa species. *Nat. Biotechnol.* **1999**, *17*, 969-973.
- (43) Morin, J. G.; Hastings, J. W. Energy transfer in a bioluminescent system. *J. Cell. Physiol.* **1971**, *77*, 313-318.
- (44) Matz, M. V.; Lukyanov, K. A.; Lukyanov, S. A. Family of the green fluorescent protein: journey to the end of the rainbow. *Bioessays* **2002**, *24*, 953-959.
- (45) Labas, Y. A.; Gurskaya, N. G.; Yanushevich, Y. G.; Fradkov, A. F.; Lukyanov, K. A.; Lukyanov, S. A.; Matz, M. V. Diversity and evolution of the green fluorescent protein family. *Proc. Natl. Acad. Sci. U. S. A.* **2002**, *99*, 4256-4261.
- (46) Shagin, D. A.; Barsova, E. V.; Yanushevich, Y. G.; Fradkov, A. F.; Lukyanov, K. A.; Labas, Y. A.; Semenova, T. N.; Ugalde, J. A.; Meyers, A.; Nunez, J. M.; Widder, E. A.; Lukyanov, S. A.; Matz, M. V. GFP-like proteins as ubiquitous metazoan superfamily: evolution of functional features and structural complexity. *Mol. Biol. Evol.* **2004**, *21*, 841-850.
- (47) Baumann, D.; Cook, M.; Ma, L.; Mushegian, A.; Sanders, E.; Schwartz, J.; Yu, C. R. A family of GFP-like proteins with different spectral properties in lancelet *Branchiostoma floridae*. *Biol. Direct* **2008**, *3*, 28.
- (48) Deheyn, D. D.; Kubokawa, K.; McCarthy, J. K.; Murakami, A.; Porrachia, M.; Rouse, G. W.; Holland, N. D. Endogenous green fluorescent protein (GFP) in amphioxus. *Biol. Bull.* **2007**, *213*.
- (49) Salih, A.; Larkum, A.; Cox, G.; Kuhl, M.; Hoegh-Guldberg, O. Fluorescent pigments in corals are photoprotective. *Nature* **2000**, *408*, 850-853.
- (50) Matz, M. V.; Marshall, N. J.; Vorobyev, M. Are corals colorful? *Photochem. Photobiol.* **2006**, *82*, 345-350.
- (51) Prasher, D. C.; Eckenrode, V. K.; Ward, W. W.; Prendergast, F. G.; Cormier, M. J. Primary structure of the *Aequorea victoria* green-fluorescent protein. *Gene* **1992**, *111*, 229-233.

- (52) Ormo, M.; Cubitt, A. B.; Kallio, K.; Gross, L. A.; Tsien, R. Y.; Remington, S. J. Crystal structure of the *Aequorea victoria* green fluorescent protein. *Science* **1996**, *273*, 1392-1395.
- (53) Yang, F.; Moss, L. G.; Phillips, G. N., Jr The molecular structure of green fluorescent protein. *Nat. Biotechnol.* **1996**, *14*, 1246-1251.
- (54) Shimomura, O. Structure of the Chromophore of *Aequorea* Green Fluorescent Protein. *FEBS Lett.* **1979**, *104*, 220-222.
- (55) Cody, C. W.; Prasher, D. C.; Westler, W. M.; Prendergast, F. G.; Ward, W. W. Chemical-Structure of the Hexapeptide Chromophore of the *Aequorea* Green-Fluorescent Protein. *Biochemistry* **1993**, *32*, 1212-1218.
- (56) Cramer, A.; Whitehorn, E. A.; Tate, E.; Stemmer, W. P. Improved green fluorescent protein by molecular evolution using DNA shuffling. *Nat. Biotechnol.* **1996**, *14*, 315-319.
- (57) Cormack, B.; Valdivia, R.; Falkow, S. FACS-optimized mutants of the green fluorescent protein (GFP). *Gene* **1996**, *173*, 33-38.
- (58) Baird, G. S.; Zacharias, D. A.; Tsien, R. Y. Circular permutation and receptor insertion within green fluorescent proteins. *Proc. Natl. Acad. Sci. U. S. A.* **1999**, *96*, 11241-11246.
- (59) Li, Y.; Sierra, A. M.; Ai, H. W.; Campbell, R. E. Identification of sites within a monomeric red fluorescent protein that tolerate peptide insertion and testing of corresponding circular permutations. *Photochem. Photobiol.* **2008**, *84*, 111-119.
- (60) Carlson, H. J.; Cotton, D. W.; Campbell, R. E. Circularly permuted monomeric red fluorescent proteins with new termini in the beta-sheet. *Protein Sci.* **2010**, *19*, 1490-1499.
- (61) Hsu, S. D.; Blaser, G.; Jackson, S. E. The folding, stability and conformational dynamics of beta-barrel fluorescent proteins. *Chem. Soc. Rev.* **2009**, *38*, 2951-2965.
- (62) Heim, R.; Prasher, D. C.; Tsien, R. Y. Wavelength Mutations and Posttranslational Autoxidation of Green Fluorescent Protein. *Proc. Natl. Acad. Sci. U. S. A.* **1994**, *91*, 12501-12504.
- (63) Rosenow, M.; Patel, H.; Wachter, R. Oxidative chemistry in the GFP active site leads to covalent cross-linking of a modified leucine side chain with a histidine imidazole: Implications for the mechanism of chromophore formation. *Biochemistry* **2005**, *44*, 8303-8311.
- (64) Katranidis, A.; Atta, D.; Schlesinger, R.; Nierhaus, K. H.; Choli-Papadopoulou, T.; Gregor, I.; Gerrits, M.; Bueldt, G.; Fitter, J. Fast Biosynthesis of GFP Molecules: A Single-Molecule Fluorescence Study. *Angew. Chem. Int. Ed.* **2009**, *48*, 1758-1761.

- (65) Wachter, R. M.; Watkins, J. L.; Kim, H. Mechanistic Diversity of Red Fluorescence Acquisition by GFP-like Proteins. *Biochemistry* **2010**, *49*, 7417-7427.
- (66) Strack, R. L.; Strongin, D. E.; Mets, L.; Glick, B. S.; Keenan, R. J. Chromophore Formation in DsRed Occurs by a Branched Pathway. *J. Am. Chem. Soc.* **2010**, *132*, 8496-8505.
- (67) Subach, O. M.; Malashkevich, V. N.; Zencheck, W. D.; Morozova, K. S.; Piatkevich, K. D.; Almo, S. C.; Verkhusha, V. V. Structural Characterization of Acylimine-Containing Blue and Red Chromophores in mTagBFP and TagRFP Fluorescent Proteins. *Chem. Biol.* **2010**, *17*, 333-341.
- (68) Karasawa, S.; Araki, T.; Nagai, T.; Mizuno, H.; Miyawaki, A. Cyan-emitting and orange-emitting fluorescent proteins as a donor/acceptor pair for fluorescence resonance energy transfer. *Biochem. J.* **2004**, *381*, 307-312.
- (69) Kikuchi, A.; Fukumura, E.; Karasawa, S.; Mizuno, H.; Miyawaki, A.; Shiro, Y. Structural characterization of a thiazoline-containing chromophore in an orange fluorescent protein, monomeric Kusabira Orange. *Biochemistry* **2008**, *47*, 11573-11580.
- (70) Prescott, M.; Ling, M.; Beddoe, T.; Oakley, A.; Dove, S.; Hoegh-Guldberg, O.; Devenish, R.; Rossjohn, J. The 2.2 Å crystal structure of a pocilloporin pigment reveals a nonplanar chromophore conformation. *Structure* **2003**, *11*, 275-284.
- (71) Petersen, J.; Wilmann, P.; Beddoe, T.; Oakley, A.; Devenish, R.; Prescott, M.; Rossjohn, J. The 2.0-angstrom crystal structure of eqFP611, a far red fluorescent protein from the sea anemone *Entacmaea quadricolor*. *J. Biol. Chem.* **2003**, *278*, 44626-44631.
- (72) Henderson, J. N.; Ai, H.; Campbell, R. E.; Remington, S. J. Structural basis for reversible photobleaching of a green fluorescent protein homologue. *Proc. Natl. Acad. Sci. U. S. A.* **2007**, *104*, 6672-6677.
- (73) Gross, L. A.; Baird, G. S.; Hoffman, R. C.; Baldrige, K. K.; Tsien, R. Y. The structure of the chromophore within DsRed, a red fluorescent protein from coral. *Proc. Natl. Acad. Sci. U. S. A.* **2000**, *97*, 11990-11995.
- (74) Remington, S. J.; Wachter, R. M.; Yarbrough, D. K.; Branchaud, B.; Anderson, D. C.; Kallio, K.; Lukyanov, K. A. zFP538, a yellow-fluorescent protein from *Zoanthus*, contains a novel three-ring chromophore. *Biochemistry* **2005**, *44*, 202-212.
- (75) Ando, R.; Hama, H.; Yamamoto-Hino, M.; Mizuno, H.; Miyawaki, A. An optical marker based on the UV-induced green-to-red

- photoconversion of a fluorescent protein. *Proc. Natl. Acad. Sci. U. S. A.* **2002**, *99*, 12651-12656.
- (76) Mizuno, H.; Mal, T.; Tong, K.; Ando, R.; Furuta, T.; Ikura, M.; Miyawakil, A. Photo-induced peptide cleavage in the green-to-red conversion of a fluorescent protein. *Mol. Cell* **2003**, *12*, 1051-1058.
- (77) Quillin, M.; Anstrom, D.; Shu, X.; O'Leary, S.; Kallio, K.; Chudakov, D.; Remington, S. Kindling fluorescent protein from *Anemonia sulcata*: Dark-state structure at 1.38 angstrom resolution. *Biochemistry* **2005**, *44*, 5774-5787.
- (78) Tretyakova, Y. A.; Pakhomov, A. A.; Martynov, V. I. Chromophore structure of the kindling fluorescent protein asFP595 from *Anemonia sulcata*. *J. Am. Chem. Soc.* **2007**, *129*, 7748-7749.
- (79) Lippincott-Schwartz, J.; Altan-Bonnet, N.; Patterson, G. Photobleaching and photoactivation: following protein dynamics in living cells. *Nat. Cell Biol.* **2003**, S7-S14.
- (80) Greenbaum, L.; Rothmann, C.; Lavie, R.; Malik, Z. Green fluorescent protein photobleaching: a model for protein damage by endogenous and exogenous singlet oxygen. *Biol. Chem.* **2000**, *381*, 1251-1258.
- (81) Alvarez, L.; Levin, C. H.; Merola, F.; Bizouarn, T.; Pasquier, H.; Baciou, L.; Rusconi, F.; Erard, M. Are the Fluorescent Properties of the Cyan Fluorescent Protein Sensitive to Conditions of Oxidative Stress? *Photochem. Photobiol.* **2010**, *86*, 55-61.
- (82) Henderson, J. N.; Gepshtein, R.; Heenan, J. R.; Kallio, K.; Huppert, D.; Remington, S. J. Structure and mechanism of the photoactivatable green fluorescent protein. *J. Am. Chem. Soc.* **2009**, *131*, 4176-4177.
- (83) Dickson, R.; Cubitt, A.; Tsien, R.; Moerner, W. On/off blinking and switching behaviour of single molecules of green fluorescent protein. *Nature* **1997**, *388*, 355-358.
- (84) Shaner, N. C.; Lin, M. Z.; McKeown, M. R.; Steinbach, P. A.; Hazelwood, K. L.; Davidson, M. W.; Tsien, R. Y. Improving the photostability of bright monomeric orange and red fluorescent proteins. *Nat. Methods* **2008**, *5*, 545-551.
- (85) Ai, H.; Hazelwood, K. L.; Davidson, M. W.; Campbell, R. E. Fluorescent protein FRET pairs for ratiometric imaging of dual biosensors. *Nat. Methods* **2008**, *5*, 401-403.
- (86) Yang, T.; Cheng, L.; Kain, S. Optimized codon usage and chromophore mutations provide enhanced sensitivity with the green fluorescent protein. *Nucleic Acids Res.* **1996**, *24*, 4592-4593.

- (87) Campbell, R.; Tour, O.; Palmer, A.; Steinbach, P.; Baird, G.; Zacharias, D.; Tsien, R. A monomeric red fluorescent protein. *Proc. Natl. Acad. Sci. U. S. A.* **2002**, *99*, 7877-7882.
- (88) Dai, M.; Fisher, H. E.; Temirov, J.; Kiss, C.; Phipps, M. E.; Pavlik, P.; Werner, J. H.; Bradbury, A. R. M. The creation of a novel fluorescent protein by guided consensus engineering. *Protein Eng. Des. Sel.* **2007**, *20*, 69-79.
- (89) Hoi, H.; Shaner, N. C.; Davidson, M. W.; Cairo, C. W.; Wang, J.; Campbell, R. E. A monomeric photoconvertible fluorescent protein for imaging of dynamic protein localization. *J. Mol. Biol.* **2010**, *401*, 776-791.
- (90) Wall, M. A.; Socolich, M.; Ranganathan, R. The structural basis for red fluorescence in the tetrameric GFP homolog DsRed. *Nat. Struct. Biol.* **2000**, *7*, 1133-1138.
- (91) Yarbrough, D.; Wachter, R. M.; Kallio, K.; Matz, M. V.; Remington, S. J. Refined crystal structure of DsRed, a red fluorescent protein from coral, at 2.0-Å resolution. *Proc. Natl. Acad. Sci. U. S. A.* **2001**, *98*, 462-467.
- (92) Ai, H. W.; Henderson, J. N.; Remington, S. J.; Campbell, R. E. Directed evolution of a monomeric, bright and photostable version of *Clavularia* cyan fluorescent protein: structural characterization and applications in fluorescence imaging. *Biochem. J.* **2006**, *400*, 531-540.
- (93) Adam, V.; Nienhaus, K.; Bourgeois, D.; Nienhaus, G. U. Structural basis of enhanced photoconversion yield in green fluorescent protein-like protein Dendra2. *Biochemistry* **2009**, *48*, 4905-4915.
- (94) Henderson, J. N.; Remington, S. J. Crystal structures and mutational analysis of amFP486, a cyan fluorescent protein from *Anemonia majano*. *Proc. Natl. Acad. Sci. U. S. A.* **2005**, *102*, 12712-12717.
- (95) Kogure, T.; Karasawa, S.; Araki, T.; Saito, K.; Kinjo, M.; Miyawaki, A. A fluorescent variant of a protein from the stony coral *Montipora* facilitates dual-color single-laser fluorescence cross-correlation spectroscopy. *Nat. Biotechnol.* **2006**, *24*, 577-581.
- (96) Karasawa, S.; Araki, T.; Yamamoto-Hino, M.; Miyawaki, A. A green-emitting fluorescent protein from *Galaxeidae* coral and its monomeric version for use in fluorescent labeling. *J. Biol. Chem.* **2003**, *278*, 34167-34171.
- (97) Gurskaya, N.; Verkhusha, V.; Shcheglov, A.; Staroverov, D.; Chepurnykh, T.; Fradkov, A.; Lukyanov, S.; Lukyanov, K. Engineering of a monomeric green-to-red photoactivatable fluorescent protein induced by blue light. *Nat. Biotechnol.* **2006**, *24*, 461-465.

- (98) Wiedenmann, J.; Ivanchenko, S.; Oswald, F.; Schmitt, F.; Rocker, C.; Salih, A.; Spindler, K.; Nienhaus, G. EosFP, a fluorescent marker protein with UV-inducible green-to-red fluorescence conversion. *Proc. Natl. Acad. Sci. U. S. A.* **2004**, *101*, 15905-15910.
- (99) Ilagan, R. P.; Rhoades, E.; Gruber, D. F.; Kao, H. T.; Pieribone, V. A.; Regan, L. A new bright green-emitting fluorescent protein--engineered monomeric and dimeric forms. *FEBS J.* **2010**, *277*, 1967-1978.
- (100) McKinney, S. A.; Murphy, C. S.; Hazelwood, K. L.; Davidson, M. W.; Looger, L. L. A bright and photostable photoconvertible fluorescent protein. *Nat. Methods* **2009**, *6*, 131-133.
- (101) Zhang, M.; Chang, H.; Zhang, Y.; Yu, J.; Wu, L.; Ji, W.; Chen, J.; Liu, B.; Lu, J.; Liu, Y.; Zhang, J.; Xu, P.; Xu, T. Rational design of true monomeric and bright photoactivatable fluorescent proteins. *Nat. Methods* **2012**, *9*, 727-729.
- (102) Lin, M. Z.; McKeown, M. R.; Ng, H.; Aguilera, T. A.; Shaner, N. C.; Campbell, R. E.; Adams, S. R.; Gross, L. A.; Ma, W.; Alber, T.; Tsien, R. Y. Autofluorescent Proteins with Excitation in the Optical Window for Intravital Imaging in Mammals. *Chem. Biol.* **2009**, *16*, 1169-1179.
- (103) Shaner, N. C.; Campbell, R. E.; Steinbach, P. A.; Giepmans, B. N.; Palmer, A. E.; Tsien, R. Y. Improved monomeric red, orange and yellow fluorescent proteins derived from *Discosoma* sp. red fluorescent protein. *Nat. Biotechnol.* **2004**, *22*, 1567-1572.
- (104) Ai, H.; Olenych, S. G.; Wong, P.; Davidson, M. W.; Campbell, R. E. Hue-shifted monomeric variants of *Clavularia* cyan fluorescent protein: identification of the molecular determinants of color and applications in fluorescence imaging. *BMC Biol.* **2008**, *6*, 13.
- (105) Zacharias, D.; Violin, J.; Newton, A.; Tsien, R. Partitioning of lipid-modified monomeric GFPs into membrane microdomains of live cells. *Science* **2002**, *296*, 913-916.
- (106) Tomosugi, W.; Matsuda, T.; Tani, T.; Nemoto, T.; Kotera, I.; Saito, K.; Horikawa, K.; Nagai, T. An ultramarine fluorescent protein with increased photostability and pH insensitivity. *Nat. Methods* **2009**, *6*, 351-353.
- (107) Ai, H.; Shaner, N. C.; Cheng, Z.; Tsien, R. Y.; Campbell, R. E. Exploration of new chromophore structures leads to the identification of improved blue fluorescent proteins. *Biochemistry* **2007**, *46*, 5904-5910.
- (108) Heim, R.; Tsien, R. Y. Engineering green fluorescent protein for improved brightness, longer wavelengths and fluorescence resonance energy transfer. *Curr. Biol.* **1996**, *6*, 178-182.

- (109) Shu, X.; Shaner, N. C.; Yarbrough, C. A.; Tsien, R. Y.; Remington, S. J. Novel chromophores and buried charges control color in mFruits. *Biochemistry* **2006**, *45*, 9639-9647.
- (110) Subach, O. M.; Gundorov, I. S.; Yoshimura, M.; Subach, F. V.; Zhang, J.; Gruenwald, D.; Souslova, E. A.; Chudakov, D. M.; Verkhusha, V. V. Conversion of red fluorescent protein into a bright blue probe. *Chem. Biol.* **2008**, *15*, 1116-1124.
- (111) Wang, L.; Jackson, W.; Steinbach, P.; Tsien, R. Evolution of new nonantibody proteins via iterative somatic hypermutation. *Proc. Natl. Acad. Sci. U. S. A.* **2004**, *101*, 16745-16749.
- (112) Shu, X.; Wang, L.; Colip, L.; Kallio, K.; Remington, S. J. Unique interactions between the chromophore and glutamate 16 lead to far-red emission in a red fluorescent protein. *Protein Sci.* **2009**, *18*, 460-466.
- (113) Wiedenmann, J.; Vallone, B.; Renzi, F.; Nienhaus, K.; Ivanchenko, S.; Rocker, C.; Nienhaus, G. Red fluorescent protein eqFP611 and its genetically engineered dimeric variants. *J. Biomed. Opt.* **2005**, *10*, 014003.
- (114) Merzlyak, E. M.; Goedhart, J.; Shcherbo, D.; Bulina, M. E.; Shcheglov, A. S.; Fradkov, A. F.; Gaintzeva, A.; Lukyanov, K. A.; Lukyanov, S.; Gadella, T. W. J.; Chudakov, D. M. Bright monomeric red fluorescent protein with an extended fluorescence lifetime. *Nat. Methods* **2007**, *4*, 555-557.
- (115) Kredel, S.; Oswald, F.; Nienhaus, K.; Deuschle, K.; Roecker, C.; Wolff, M.; Heilker, R.; Nienhaus, G. U.; Wiedenmann, J. mRuby, a Bright Monomeric Red Fluorescent Protein for Labeling of Subcellular Structures. *PLoS One* **2009**, *4*, e4391.
- (116) Shcherbo, D.; Murphy, C. S.; Ermakova, G. V.; Solovieva, E. A.; Chepurnykh, T. V.; Shcheglov, A. S.; Verkhusha, V. V.; Pletnev, V. Z.; Hazelwood, K. L.; Roche, P. M.; Lukyanov, S.; Zaraisky, A. G.; Davidson, M. W.; Chudakov, D. M. Far-red fluorescent tags for protein imaging in living tissues. *Biochem. J.* **2009**, *418*, 567-574.
- (117) Shcherbo, D.; Shemiakina, I. I.; Ryabova, A. V.; Luker, K. E.; Schmidt, B. T.; Souslova, E. A.; Gorodnicheva, T. V.; Strukova, L.; Shidlovskiy, K. M.; Britanova, O. V.; Zaraisky, A. G.; Lukyanov, K. A.; Loschenov, V. B.; Luker, G. D.; Chudakov, D. M. Near-infrared fluorescent proteins. *Nat. Methods* **2010**, *7*, 827-U1520.
- (118) Subach, F. V.; Piatkevich, K. D.; Verkhusha, V. V. Directed molecular evolution to design advanced red fluorescent proteins. *Nat. Methods* **2011**, *8*, 1019-1026.
- (119) Goedhart, J.; van Weeren, L.; Hink, M. A.; Vischer, N. O. E.; Jalink, K.; Gadella, T. W. J., Jr. Bright cyan fluorescent protein variants

- identified by fluorescence lifetime screening. *Nat. Methods* **2010**, *7*, 137-U74.
- (120) Markwardt, M. L.; Kremers, G.; Kraft, C. A.; Ray, K.; Cranfill, P. J. C.; Wilson, K. A.; Day, R. N.; Wachter, R. M.; Davidson, M. W.; Rizzo, M. A. An Improved Cerulean Fluorescent Protein with Enhanced Brightness and Reduced Reversible Photoswitching. *PLoS One* **2011**, *6*, e17896.
- (121) Zapata-Hommer, O.; Griesbeck, O. Efficiently folding and circularly permuted variants of the Sapphire mutant of GFP. *BMC Biotechnol.* **2003**, *3*, 5.
- (122) Shcherbakova, D. M.; Hink, M. A.; Joosen, L.; Gadella, T. W. J.; Verkhusha, V. V. An Orange Fluorescent Protein with a Large Stokes Shift for Single-Excitation Multicolor FCCS and FRET Imaging. *J. Am. Chem. Soc.* **2012**, *134*, 7913-7923.
- (123) Carlson, H. J.; Campbell, R. E. Genetically encoded FRET-based biosensors for multiparameter fluorescence imaging. *Curr. Opin. Biotechnol.* **2009**, *20*, 19-27.
- (124) Ganesan, S.; Ameer-beg, S.; Ng, T.; Vojnovic, B.; Wouters, F. A dark yellow fluorescent protein (YFP)-based Resonance Energy-Accepting Chromoprotein (REACH) for Forster resonance energy transfer with GFP. *Proc. Natl. Acad. Sci. U. S. A.* **2006**, *103*, 4089-4094.
- (125) Fitzpatrick, J. A.; Lillemeier, B. F. Fluorescence correlation spectroscopy: linking molecular dynamics to biological function in vitro and in situ. *Curr. Opin. Struct. Biol.* **2011**, *21*, 650-660.
- (126) Lukyanov, K.; Chudakov, D.; Lukyanov, S.; Verkhusha, V. Photoactivatable fluorescent proteins. *Nat. Rev. Mol. Cell Biol.* **2005**, *6*, 885-891.
- (127) Chudakov, D. M.; Lukyanov, S.; Lukyanov, K. A. Tracking intracellular protein movements using photoswitchable fluorescent proteins PS-CFP2 and Dendra2. *Nat. Protoc.* **2007**, *2*, 2024-2032.
- (128) Patterson, G. H. Highlights of the optical highlighter fluorescent proteins. *J. Microsc.* **2011**, *243*, 1-7.
- (129) Ghosh, I.; Hamilton, A.; Regan, L. Antiparallel leucine zipper-directed protein reassembly: Application to the green fluorescent protein. *J. Am. Chem. Soc.* **2000**, *122*, 5658-5659.
- (130) Shyu, Y. J.; Hu, C. D. Fluorescence complementation: an emerging tool for biological research. *Trends Biotechnol.* **2008**, *26*, 622-630.
- (131) Chu, J.; Zhang, Z.; Zheng, Y.; Yang, J.; Qin, L.; Lu, J.; Huang, Z. L.; Zeng, S.; Luo, Q. A novel far-red bimolecular fluorescence complementation system that allows for efficient visualization of protein

- interactions under physiological conditions. *Biosens. Bioelectron.* **2009**, *25*, 234-239.
- (132) Magliery, T. J.; Wilson, C. G.; Pan, W.; Mishler, D.; Ghosh, I.; Hamilton, A. D.; Regan, L. Detecting protein-protein interactions with a green fluorescent protein fragment reassembly trap: scope and mechanism. *J. Am. Chem. Soc.* **2005**, *127*, 146-157.
- (133) Alford, S. C.; Abdelfattah, A. S.; Ding, Y.; Campbell, R. E. A fluorogenic red fluorescent protein heterodimer. *Chem. Biol.* **2012**, *19*, 353-360.
- (134) Valeur, B. In *Non-radiative energy transfer; Molecular Fluorescence: Principles and Applications*; Wiley-VCH: Verlag GmbH, **2001**; pp 113-122.
- (135) Forster, T. Transfer mechanisms of electronic excitation. *Discuss. Faraday Soc.* **1959**, *27*, 7-17.
- (136) Frommer, W. B.; Davidson, M. W.; Campbell, R. E. Genetically encoded biosensors based on engineered fluorescent proteins. *Chem. Soc. Rev.* **2009**, *38*, 2833-2841.
- (137) Miyawaki, A.; Llopis, J.; Heim, R.; McCaffery, J. M.; Adams, J. A.; Ikura, M.; Tsien, R. Y. Fluorescent indicators for Ca²⁺ based on green fluorescent proteins and calmodulin. *Nature* **1997**, *388*, 882-887.
- (138) Ding, Y.; Ai, H. W.; Hoi, H.; Campbell, R. E. Forster resonance energy transfer-based biosensors for multiparameter ratiometric imaging of Ca²⁺ dynamics and caspase-3 activity in single cells. *Anal. Chem.* **2011**, *83*, 9687-9693.
- (139) Horikawa, K.; Yamada, Y.; Matsuda, T.; Kobayashi, K.; Hashimoto, M.; Matsu-ura, T.; Miyawaki, A.; Michikawa, T.; Mikoshiba, K.; Nagai, T. Spontaneous network activity visualized by ultrasensitive Ca(2+) indicators, yellow Cameleon-Nano. *Nat. Methods* **2010**, *7*, 729-732.
- (140) Dittmer, P. J.; Miranda, J. G.; Gorski, J. A.; Palmer, A. E. Genetically encoded sensors to elucidate spatial distribution of cellular zinc. *J. Biol. Chem.* **2009**, *284*, 16289-16297.
- (141) DiPilato, L. M.; Cheng, X.; Zhang, J. Fluorescent indicators of cAMP and Epac activation reveal differential dynamics of cAMP signaling within discrete subcellular compartments. *Proc. Natl. Acad. Sci. U. S. A.* **2004**, *101*, 16513-16518.
- (142) Lager, I.; Looger, L. L.; Hilpert, M.; Lalonde, S.; Frommer, W. B. Conversion of a putative Agrobacterium sugar-binding protein into a FRET sensor with high selectivity for sucrose. *J. Biol. Chem.* **2006**, *281*, 30875-30883.

- (143) Allen, M. D.; Zhang, J. Subcellular dynamics of protein kinase A activity visualized by FRET-based reporters. *Biochem. Biophys. Res. Commun.* **2006**, *348*, 716-721.
- (144) Zhang, J.; Hupfeld, C. J.; Taylor, S. S.; Olefsky, J. M.; Tsien, R. Y. Insulin disrupts beta-adrenergic signalling to protein kinase A in adipocytes. *Nature* **2005**, *437*, 569-573.
- (145) Zhang, J.; Ma, Y.; Taylor, S. S.; Tsien, R. Y. Genetically encoded reporters of protein kinase A activity reveal impact of substrate tethering. *Proc. Natl. Acad. Sci. U. S. A.* **2001**, *98*, 14997-15002.
- (146) Kunkel, M. T.; Ni, Q.; Tsien, R. Y.; Zhang, J.; Newton, A. C. Spatio-temporal dynamics of protein kinase B/Akt signaling revealed by a genetically encoded fluorescent reporter. *J. Biol. Chem.* **2005**, *280*, 5581-5587.
- (147) Wang, Y.; Botvinick, E. L.; Zhao, Y.; Berns, M. W.; Usami, S.; Tsien, R. Y.; Chien, S. Visualizing the mechanical activation of Src. *Nature* **2005**, *434*, 1040-1045.
- (148) Hwang, Y. C.; Chu, J. J.; Yang, P. L.; Chen, W.; Yates, M. V. Rapid identification of inhibitors that interfere with poliovirus replication using a cell-based assay. *Antiviral Res.* **2008**, *77*, 232-236.
- (149) Sabariego, R.; Picazo, F.; Domingo, B.; Franco, S.; Martinez, M.; Llopis, J. Fluorescence Resonance Energy Transfer-Based Assay for Characterization of Hepatitis C Virus NS3-4A Protease Activity in Live Cells. *Antimicrobial Agents Chemother.* **2009**, *53*, 728-734.
- (150) Zhu, X.; Fu, A.; Luo, K. Q. A high-throughput fluorescence resonance energy transfer (FRET)-based endothelial cell apoptosis assay and its application for screening vascular disrupting agents. *Biochem. Biophys. Res. Commun.* **2012**, *418*, 641-646.
- (151) Lam, A. J.; St-Pierre, F.; Gong, Y.; Marshall, J. D.; Cranfill, P. J.; Baird, M. A.; McKeown, M. R.; Wiedenmann, J.; Davidson, M. W.; Schnitzer, M. J.; Tsien, R. Y.; Lin, M. Z. Improving FRET dynamic range with bright green and red fluorescent proteins. *Nat. Methods* **2012**, *9*, 1005-1012.
- (152) Ouyang, M.; Huang, H.; Shaner, N. C.; Remacle, A. G.; Shiryaev, S. A.; Strongin, A. Y.; Tsien, R. Y.; Wang, Y. Simultaneous visualization of protumorigenic Src and MT1-MMP activities with fluorescence resonance energy transfer. *Cancer Res.* **2010**, *70*, 2204-2212.
- (153) Grant, D. M.; Zhang, W.; McGhee, E. J.; Bunney, T. D.; Talbot, C. B.; Kumar, S.; Munro, I.; Dunsby, C.; Neil, M. A.; Katan, M.; French, P. M. Multiplexed FRET to image multiple signaling events in live cells. *Biophys. J.* **2008**, *95*, L69-71.

- (154) Komatsu, N.; Aoki, K.; Yamada, M.; Yukinaga, H.; Fujita, Y.; Kamioka, Y.; Matsuda, M. Development of an optimized backbone of FRET biosensors for kinases and GTPases. *Mol. Biol. Cell* **2011**, *22*, 4647-4656.
- (155) Wachter, R. M.; Remington, S. J. Sensitivity of the yellow variant of green fluorescent protein to halides and nitrate. *Curr. Biol.* **1999**, *9*, R628-9.
- (156) Jayaraman, S.; Haggie, P.; Wachter, R. M.; Remington, S. J.; Verkman, A. S. Mechanism and cellular applications of a green fluorescent protein-based halide sensor. *J. Biol. Chem.* **2000**, *275*, 6047-6050.
- (157) Llopis, J.; McCaffery, J. M.; Miyawaki, A.; Farquhar, M. G.; Tsien, R. Y. Measurement of cytosolic, mitochondrial, and Golgi pH in single living cells with green fluorescent proteins. *Proc. Natl. Acad. Sci. U. S. A.* **1998**, *95*, 6803-6808.
- (158) Miesenbock, G.; De Angelis, D.; Rothman, J. Visualizing secretion and synaptic transmission with pH-sensitive green fluorescent proteins. *Nature* **1998**, *394*, 192-195.
- (159) Wachter, R. M.; Yarbrough, D.; Kallio, K.; Remington, S. J. Crystallographic and energetic analysis of binding of selected anions to the yellow variants of green fluorescent protein. *J. Mol. Biol.* **2000**, *301*, 157-171.
- (160) Galletta, L. J.; Haggie, P. M.; Verkman, A. S. Green fluorescent protein-based halide indicators with improved chloride and iodide affinities. *FEBS Lett.* **2001**, *499*, 220-224.
- (161) Rhoden, K. J.; Cianchetta, S.; Duchi, S.; Romeo, G. Fluorescence quantitation of thyrocyte iodide accumulation with the yellow fluorescent protein variant YFP-H148Q/I152L. *Anal. Biochem.* **2008**, *373*, 239-246.
- (162) Hanson, G. T.; Aggeler, R.; Oglesbee, D.; Cannon, M.; Capaldi, R. A.; Tsien, R. Y.; Remington, S. J. Investigating mitochondrial redox potential with redox-sensitive green fluorescent protein indicators. *J. Biol. Chem.* **2004**, *279*, 13044-13053.
- (163) Ostergaard, H.; Henriksen, A.; Hansen, F. G.; Winther, J. R. Shedding light on disulfide bond formation: engineering a redox switch in green fluorescent protein. *EMBO J.* **2001**, *20*, 5853-5862.
- (164) Nagai, T.; Sawano, A.; Park, E. S.; Miyawaki, A. Circularly permuted green fluorescent proteins engineered to sense Ca²⁺. *Proc. Natl. Acad. Sci. U. S. A.* **2001**, *98*, 3197-3202.

- (165) Nakai, J.; Ohkura, M.; Imoto, K. A high signal-to-noise Ca²⁺ probe composed of a single green fluorescent protein. *Nat. Biotechnol.* **2001**, *19*, 137-141.
- (166) Wang, Q.; Shui, B.; Kotlikoff, M. I.; Sondermann, H. Structural basis for calcium sensing by GCaMP2. *Structure* **2008**, *16*, 1817-1827.
- (167) Akerboom, J.; Rivera, J. D.; Guilbe, M. M.; Malave, E. C.; Hernandez, H. H.; Tian, L.; Hires, S. A.; Marvin, J. S.; Looger, L. L.; Schreier, E. R. Crystal structures of the GCaMP calcium sensor reveal the mechanism of fluorescence signal change and aid rational design. *J. Biol. Chem.* **2009**, *284*, 6455-6464.
- (168) Zhao, Y.; Araki, S.; Wu, J.; Teramoto, T.; Chang, Y. F.; Nakano, M.; Abdelfattah, A. S.; Fujiwara, M.; Ishihara, T.; Nagai, T.; Campbell, R. E. An expanded palette of genetically encoded Ca²⁺ indicators. *Science* **2011**, *333*, 1888-1891.
- (169) McEvoy, A. L.; Hoi, H.; Bates, M.; Platonova, E.; Cranfill, P. J.; Baird, M. A.; Davidson, M. W.; Ewers, H.; Liphardt, J.; Campbell, R. E. mMaple: a photoconvertible fluorescent protein for use in multiple imaging modalities. *PLoS one* **2012**, *7*, e51314.
- (170) Tsutsui, H.; Karasawa, S.; Shimizu, H.; Nukina, N.; Miyawaki, A. Semi-rational engineering of a coral fluorescent protein into an efficient highlighter. *EMBO Rep.* **2005**, *6*, 233-238.
- (171) Chatteraj, M.; King, B. A.; Bublitz, G. U.; Boxer, S. G. Ultra-fast excited state dynamics in green fluorescent protein: multiple states and proton transfer. *Proc. Natl. Acad. Sci. U. S. A.* **1996**, *93*, 8362-8367.
- (172) Tsutsui, H.; Shimizu, H.; Mizuno, H.; Nukina, N.; Furuta, T.; Miyawaki, A. The E1 Mechanism in Photo-Induced beta-Elimination Reactions for Green-to-Red Conversion of Fluorescent Proteins. *Chem. Biol.* **2009**, *16*, 1140-1147.
- (173) Chudakov, D.; Verkhusha, V.; Staroverov, D.; Souslova, E.; Lukyanov, S.; Lukyanov, K. Photoswitchable cyan fluorescent protein for protein tracking. *Nat. Biotechnol.* **2004**, *22*, 1435-1439.
- (174) Kremers, G.; Hazelwood, K. L.; Murphy, C. S.; Davidson, M. W.; Piston, D. W. Photoconversion in orange and red fluorescent proteins. *Nat. Methods* **2009**, *6*, 355-U55.
- (175) Patterson, G. H.; Lippincott-Schwartz, J. A photoactivatable GFP for selective photolabeling of proteins and cells. *Science* **2002**, *297*, 1873-1877.
- (176) Ando, R.; Mizuno, H.; Miyawaki, A. Regulated fast nucleocytoplasmic shuttling observed by reversible protein highlighting. *Science* **2004**, *306*, 1370-1373.

- (177) Patterson, G.; Davidson, M.; Manley, S.; Lippincott-Schwartz, J. Superresolution imaging using single-molecule localization. *Annu. Rev. Phys. Chem.* **2010**, *61*, 345-367.
- (178) Patterson, G.; Davidson, M.; Manley, S.; Lippincott-Schwartz, J. Superresolution Imaging using Single-Molecule Localization. *Annu. Rev. Phys. Chem.* **2010**, *61*, 345-367.
- (179) Hayashi, I.; Mizuno, H.; Tong, K. I.; Furuta, T.; Tanaka, F.; Yoshimura, M.; Miyawaki, A.; Ikura, M. Crystallographic evidence for water-assisted photo-induced peptide cleavage in the stony coral fluorescent protein kaede. *J. Mol. Biol.* **2007**, *372*, 918-926.
- (180) Habuchi, S.; Tsutsui, H.; Kochaniak, A. B.; Miyawaki, A.; van Oijen, A. M. mKikGR, a Monomeric Photoswitchable Fluorescent Protein. *PLoS One* **2008**, *3*, e3944.
- (181) Nienhaus, G.; Nienhaus, K.; Holzle, A.; Ivanchenko, S.; Renzi, F.; Oswald, F.; Wolff, M.; Schmitt, F.; Rocker, C.; Vallone, B.; Weidemann, W.; Heilker, R.; Nar, H.; Wiedenmann, J. Photoconvertible fluorescent protein EosFP: Biophysical properties and cell biology applications. *Photochem. Photobiol.* **2006**, *82*, 351-358.
- (182) Fromant, M.; Blanquet, S.; Plateau, P. Direct random mutagenesis of gene-sized DNA fragments using polymerase chain reaction. *Anal. Biochem.* **1995**, *224*, 347-353.
- (183) Zhao, H.; Zha, W. In vitro 'sexual' evolution through the PCR-based staggered extension process (StEP). *Nat. Protoc.* **2006**, *1*, 1865-1871.
- (184) Dittrich, P. S.; Schafer, S. P.; Schwille, P. Characterization of the photoconversion on reaction of the fluorescent protein Kaede on the single-molecule level. *Biophys. J.* **2005**, *89*, 3446-3455.
- (185) Adam, V.; Lelimosin, M.; Boehme, S.; Desfonds, G.; Nienhaus, K.; Field, M. J.; Wiedenmann, J.; McSweeney, S.; Nienhaus, G. U.; Bourgeois, D. Structural characterization of IrisFP, an optical highlighter undergoing multiple photo-induced transformations. *Proc. Natl. Acad. Sci. U. S. A.* **2008**, *105*, 18343-18348.
- (186) Jung, G.; Mais, S.; Zumbusch, A.; Brauchle, C. The role of dark states in the photodynamics of the green fluorescent protein examined with two-color fluorescence excitation spectroscopy. *J. Phys. Chem. A* **2000**, *104*, 873-877.
- (187) Yang, L.; Kowalski, J. R.; Yacono, P.; Bajmoczy, M.; Shaw, S. K.; Froio, R. M.; Golan, D. E.; Thomas, S. M.; Luscinskas, F. W. Endothelial cell cortactin coordinates intercellular adhesion molecule-1 clustering and actin cytoskeleton remodeling during polymorphonuclear leukocyte adhesion and transmigration. *J. Immunol.* **2006**, *177*, 6440-6449.

- (188) Ho, S. N.; Hunt, H. D.; Horton, R. M.; Pullen, J. K.; Pease, L. R. Site-directed mutagenesis by overlap extension using the polymerase chain reaction. *Gene* **1989**, *77*, 51-59.
- (189) Ward, W. W. Biochemical and Physical Properties of Green Fluorescent Protein. *Green Fluorescent Protein: Properties, Applications, and Protocols, 2nd Edition* **2006**, *47*, 39-65.
- (190) Brannon, J.; Magde, D. Absolute Quantum Yield Determination by Thermal Blooming - Fluorescein. *J. Phys. Chem.* **1978**, *82*, 705-709.
- (191) Fischer, M.; Georges, J. Fluorescence quantum yield of rhodamine 6G in ethanol as a function of concentration using thermal lens spectrometry. *Chem. Phys. Lett.* **1996**, *260*, 115-118.
- (192) Staunton, D. E.; Marlin, S. D.; Stratowa, C.; Dustin, M. L.; Springer, T. A. Primary structure of ICAM-1 demonstrates interaction between members of the immunoglobulin and integrin supergene families. *Cell* **1988**, *52*, 925-933.
- (193) Wulfigg, C.; Sjaastad, M.; Davis, M. Visualizing the dynamics of T cell activation: Intracellular adhesion molecule 1 migrates rapidly to the T cell/B cell interface and acts to sustain calcium levels. *Proc. Natl. Acad. Sci. U. S. A.* **1998**, *95*, 6302-6307.
- (194) Schneider, C. A.; Rasband, W. S.; Eliceiri, K. W. NIH Image to ImageJ: 25 years of image analysis. *Nat. Methods* **2012**, *9*, 671-675.
- (195) Betzig, E.; Patterson, G. H.; Sougrat, R.; Lindwasser, O. W.; Olenych, S.; Bonifacino, J. S.; Davidson, M. W.; Lippincott-Schwartz, J.; Hess, H. F. Imaging intracellular fluorescent proteins at nanometer resolution. *Science* **2006**, *313*, 1642-1645.
- (196) Rust, M. J.; Bates, M.; Zhuang, X. Sub-diffraction-limit imaging by stochastic optical reconstruction microscopy (STORM). *Nat. Methods* **2006**, *3*, 793-795.
- (197) Hess, S. T.; Girirajan, T. P.; Mason, M. D. Ultra-high resolution imaging by fluorescence photoactivation localization microscopy. *Biophys. J.* **2006**, *91*, 4258-4272.
- (198) Abbe, E. Beiträge zur Theorie des Mikroskops und der mikroskopischen Wahrnehmung. *Archiv für Mikroskopische Anatomie* **1873**, *9*, 413-418.
- (199) Jones, S. A.; Shim, S. H.; He, J.; Zhuang, X. Fast, three-dimensional super-resolution imaging of live cells. *Nat. Methods* **2011**, *8*, 499-508.
- (200) Shroff, H.; Galbraith, C. G.; Galbraith, J. A.; Betzig, E. Live-cell photoactivated localization microscopy of nanoscale adhesion dynamics. *Nat. Methods* **2008**, *5*, 417-423.

- (201) Hell, S. W.; Wichmann, J. Breaking the diffraction resolution limit by stimulated emission: stimulated-emission-depletion fluorescence microscopy. *Opt. Lett.* **1994**, *19*, 780-782.
- (202) Klar, T. A.; Jakobs, S.; Dyba, M.; Egner, A.; Hell, S. W. Fluorescence microscopy with diffraction resolution barrier broken by stimulated emission. *Proc. Natl. Acad. Sci. U. S. A.* **2000**, *97*, 8206-8210.
- (203) Gustafsson, M. G. Surpassing the lateral resolution limit by a factor of two using structured illumination microscopy. *J. Microsc.* **2000**, *198*, 82-87.
- (204) Gustafsson, M. G. Nonlinear structured-illumination microscopy: wide-field fluorescence imaging with theoretically unlimited resolution. *Proc. Natl. Acad. Sci. U. S. A.* **2005**, *102*, 13081-13086.
- (205) Willig, K. I.; Kellner, R. R.; Medda, R.; Hein, B.; Jakobs, S.; Hell, S. W. Nanoscale resolution in GFP-based microscopy. *Nat. Methods* **2006**, *3*, 721-723.
- (206) Huang, B.; Bates, M.; Zhuang, X. Super-Resolution Fluorescence Microscopy. *Annu. Rev. Biochem.* **2009**, *78*, 993-1016.
- (207) Li, M.; Hazelbauer, G. Cellular stoichiometry of the components of the chemotaxis signaling complex. *J. Bacteriol.* **2004**, *186*, 3687-3694.
- (208) Thompson, R. E.; Larson, D. R.; Webb, W. W. Precise nanometer localization analysis for individual fluorescent probes. *Biophys. J.* **2002**, *82*, 2775-2783.
- (209) Annibale, P.; Scarselli, M.; Kodyan, A.; Radenovic, A. Photoactivatable Fluorescent Protein mEos2 Displays Repeated Photoactivation after a Long-Lived Dark State in the Red Photoconverted Form. *J. Phys. Chem. Lett.* **2010**, *1*, 1506-1510.
- (210) Greenfield, D.; McEvoy, A. L.; Shroff, H.; Crooks, G. E.; Wingreen, N. S.; Betzig, E.; Liphardt, J. Self-organization of the Escherichia coli chemotaxis network imaged with super-resolution light microscopy. *PLoS Biol.* **2009**, *7*, e1000137.
- (211) Annibale, P.; Vanni, S.; Scarselli, M.; Rothlisberger, U.; Radenovic, A. Identification of clustering artifacts in photoactivated localization microscopy. *Nat. Methods* **2011**, *8*, 527-528.
- (212) Berridge, M. J.; Lipp, P.; Bootman, M. D. The versatility and universality of calcium signalling. *Nat. Rev. Mol. Cell Biol.* **2000**, *1*, 11-21.
- (213) Russell, J. T. Imaging calcium signals in vivo: a powerful tool in physiology and pharmacology. *Br. J. Pharmacol.* **2011**, *163*, 1605-1625.

- (214) Adams, S. R. How calcium indicators work. *Cold Spring Harb Protoc.* **2010**, 2010, pdb.top70.
- (215) Kotlikoff, M. I. Genetically encoded Ca²⁺ indicators: using genetics and molecular design to understand complex physiology. *J. Physiol.* **2007**, 578, 55-67.
- (216) Palmer, A. E.; Tsien, R. Y. Measuring calcium signaling using genetically targetable fluorescent indicators. *Nat. Protoc.* **2006**, 1, 1057-1065.
- (217) Tallini, Y. N.; Ohkura, M.; Choi, B. R.; Ji, G.; Imoto, K.; Doran, R.; Lee, J.; Plan, P.; Wilson, J.; Xin, H. B.; Sanbe, A.; Gulick, J.; Mathai, J.; Robbins, J.; Salama, G.; Nakai, J.; Kotlikoff, M. I. Imaging cellular signals in the heart in vivo: Cardiac expression of the high-signal Ca²⁺ indicator GCaMP2. *Proc. Natl. Acad. Sci. U. S. A.* **2006**, 103, 4753-4758.
- (218) Tian, L.; Hires, S. A.; Mao, T.; Huber, D.; Chiappe, M. E.; Chalasani, S. H.; Petreanu, L.; Akerboom, J.; McKinney, S. A.; Schreiter, E. R.; Bargmann, C. I.; Jayaraman, V.; Svoboda, K.; Looger, L. L. Imaging neural activity in worms, flies and mice with improved GCaMP calcium indicators. *Nat. Methods* **2009**, 6, 875-881.
- (219) Subach, F. V.; Malashkevich, V. N.; Zencheck, W. D.; Xiao, H.; Filonov, G. S.; Almo, S. C.; Verkhusha, V. V. Photoactivation mechanism of PAmCherry based on crystal structures of the protein in the dark and fluorescent states. *Proc. Natl. Acad. Sci. U. S. A.* **2009**, 106, 21097-21102.
- (220) Subach, O. M.; Patterson, G. H.; Ting, L. M.; Wang, Y.; Condeelis, J. S.; Verkhusha, V. V. A photoswitchable orange-to-far-red fluorescent protein, PSmOrange. *Nat. Methods* **2011**, 8, 771-777.
- (221) Cubitt, A. B.; Heim, R.; Adams, S. R.; Boyd, A. E.; Gross, L. A.; Tsien, R. Y. Understanding, Improving and using Green Fluorescent Proteins. *Trends Biochem. Sci.* **1995**, 20.
- (222) van Thor, J. J.; Gensch, T.; Hellingwerf, K. J.; Johnson, L. N. Phototransformation of green fluorescent protein with UV and visible light leads to decarboxylation of glutamate 222. *Nat. Struct. Biol.* **2002**, 9, 37-41.
- (223) van Thor, J. J. Photoreactions and dynamics of the green fluorescent protein. *Chem. Soc. Rev.* **2009**, 38, 2935-2950.
- (224) Bizzarri, R.; Serresi, M.; Cardarelli, F.; Abbruzzetti, S.; Campanini, B.; Viappiani, C.; Beltram, F. Single amino acid replacement makes *Aequorea victoria* fluorescent proteins reversibly photoswitchable. *J. Am. Chem. Soc.* **2010**, 132, 85-95.

- (225) Subach, F. V.; Patterson, G. H.; Manley, S.; Gillette, J. M.; Lippincott-Schwartz, J.; Verkhusha, V. V. Photoactivatable mCherry for high-resolution two-color fluorescence microscopy. *Nat. Methods* **2009**, *6*, 153-159.
- (226) Subach, O. M.; Entenberg, D.; Condeelis, J. S.; Verkhusha, V. V. A FRET-Facilitated Photoswitching Using an Orange Fluorescent Protein with the Fast Photoconversion Kinetics. *J. Am. Chem. Soc.* **2012**, *134*.
- (227) Nienhaus, K.; Nienhaus, G.; Wiedenmann, J.; Nar, H. Structural basis for photo-induced protein cleavage and green-to-red conversion of fluorescent protein EosFP. *Proc. Natl. Acad. Sci. U. S. A.* **2005**, *102*, 9156-9159.
- (228) Miyawaki, A.; Griesbeck, O.; Heim, R.; Tsien, R. Y. Dynamic and quantitative Ca²⁺ measurements using improved cameleons. *Proc. Natl. Acad. Sci. U. S. A.* **1999**, *96*, 2135-2140.
- (229) Griesbeck, O.; Baird, G. S.; Campbell, R. E.; Zacharias, D. A.; Tsien, R. Y. Reducing the environmental sensitivity of yellow fluorescent protein. Mechanism and applications. *J. Biol. Chem.* **2001**, *276*, 29188-29194.
- (230) Nagai, T.; Ibata, K.; Park, E. S.; Kubota, M.; Mikoshiba, K.; Miyawaki, A. A variant of yellow fluorescent protein with fast and efficient maturation for cell-biological applications. *Nat. Biotechnol.* **2002**, *20*, 87-90.
- (231) Aglyamova, G. V.; Hunt, M. E.; Modi, C. K.; Matz, M. V. Multi-colored homologs of the green fluorescent protein from hydromedusa *Obelia* sp. *Photochem. Photobiol. Sci.* **2011**, *10*, 1303-1309.
- (232) Pakhomov, A. A.; Martynov, V. I. Probing the structural determinants of yellow fluorescence of a protein from *Phialidium* sp. *Biochem. Biophys. Res. Commun.* **2011**, *407*, 230-235.
- (233) Zaganichny, V. E.; Rudenko, N. V.; Gorokhovatsky, A. Y.; Zakharov, M. V.; Shenkarev, Z. O.; Balashova, T. A.; Arseniev, A. S. zFP538, a yellow fluorescent protein from coral, belongs to the DsRed subfamily of GFP-like proteins but possesses the unexpected site of fragmentation. *Biochemistry* **2004**, *43*, 4764-4772.
- (234) Yanushevich, Y. G.; Staroverov, D. B.; Savitsky, A. P.; Fradkov, A. F.; Gurskaya, N. G.; Bulina, M. E.; Lukyanov, K. A.; Lukyanov, S. A. A strategy for the generation of non-aggregating mutants of Anthozoa fluorescent proteins. *FEBS Lett.* **2002**, *511*, 11-14.
- (235) Gurskaya, N. G.; Savitsky, A. P.; Yanushevich, Y. G.; Lukyanov, S. A.; Lukyanov, K. A. Color transitions in coral's fluorescent proteins by site-directed mutagenesis. *BMC Biochem.* **2001**, *2*, 6.

- (236) Subach, F. V.; Verkhusha, V. V. Chromophore transformations in red fluorescent proteins. *Chem. Rev.* **2012**, *112*, 4308-4327.
- (237) Pletneva, N. V.; Pletnev, S. V.; Chudakov, D. M.; Tikhonova, T. V.; Popov, V. O.; Martynov, V. I.; Wlodawer, A.; Dauter, Z.; Pletnev, V. Z. Three-dimensional structure of yellow fluorescent protein zYFP538 from *Zoanthus* sp. at the resolution 1.8 angstrom. *Bioorg. Khim.* **2007**, *33*, 421-430.
- (238) Costantini, L. M.; Fossati, M.; Francolini, M.; Snapp, E. L. Assessing the tendency of fluorescent proteins to oligomerize under physiologic conditions. *Traffic* **2012**, *13*, 643-649.
- (239) Snapp, E. L.; Hegde, R. S.; Francolini, M.; Lombardo, F.; Colombo, S.; Pedrazzini, E.; Borgese, N.; Lippincott-Schwartz, J. Formation of stacked ER cisternae by low affinity protein interactions. *J. Cell Biol.* **2003**, *163*, 257-269.
- (240) Demuro, A.; Penna, A.; Safrina, O.; Yeromin, A. V.; Amcheslavsky, A.; Cahalan, M. D.; Parker, I. Subunit stoichiometry of human Orai1 and Orai3 channels in closed and open states. *Proc. Natl. Acad. Sci. U. S. A.* **2011**, *108*, 17832-17837.
- (241) Palmer, A. E.; Qin, Y.; Park, J. G.; McCombs, J. E. Design and application of genetically encoded biosensors. *Trends Biotechnol.* **2011**, *29*, 144-152.
- (242) Adam, V.; Moeyaert, B.; David, C. C.; Mizuno, H.; Lelimosin, M.; Dedecker, P.; Ando, R.; Miyawaki, A.; Michiels, J.; Engelborghs, Y.; Hofkens, J. Rational design of photoconvertible and biphotocchromic fluorescent proteins for advanced microscopy applications. *Chem. Biol.* **2011**, *18*, 1241-1251.
- (243) Fuchs, J.; Bohme, S.; Oswald, F.; Hedde, P. N.; Krause, M.; Wiedenmann, J.; Nienhaus, G. U. A photoactivatable marker protein for pulse-chase imaging with superresolution. *Nat. Methods* **2010**, *7*, 627-630.
- (244) Kirber, M. T.; Chen, K.; Keaney, J. F., Jr YFP photoconversion revisited: confirmation of the CFP-like species. *Nat. Methods* **2007**, *4*, 767-768.
- (245) Valentin, G.; Verheggen, C.; Pilot, T.; Neel, H.; Coppey-Moisan, M.; Bertrand, E. Photoconversion of YFP into a CFP-like species during acceptor photobleaching FRET experiments. *Nat. Methods* **2005**, *2*, 801.
- (246) Mueller, F.; Morisaki, T.; Mazza, D.; McNally, J. G. Minimizing the impact of photoswitching of fluorescent proteins on FRAP analysis. *Biophys. J.* **2012**, *102*, 1656-1665.

- (247) Heinze, K. G.; Costantino, S.; De Koninck, P.; Wiseman, P. W. Beyond photobleaching, laser illumination unbinds fluorescent proteins. *J. Phys. Chem. B* **2009**, *113*, 5225-5233.
- (248) Lubbeck, J. L.; Dean, K. M.; Ma, H.; Palmer, A. E.; Jimenez, R. Microfluidic flow cytometer for quantifying photobleaching of fluorescent proteins in cells. *Anal. Chem.* **2012**, *84*, 3929-3937.
- (249) Malkani, N.; Schmid, J. A. Some secrets of fluorescent proteins: distinct bleaching in various mounting fluids and photoactivation of cyan fluorescent proteins at YFP-excitation. *PLoS One* **2011**, *6*, e18586.
- (250) Foo, Y. H.; Naredi-Rainer, N.; Lamb, D. C.; Ahmed, S.; Wohland, T. Factors affecting the quantification of biomolecular interactions by fluorescence cross-correlation spectroscopy. *Biophys. J.* **2012**, *102*, 1174-1183.
- (251) Strack, R. L.; Strongin, D. E.; Bhattacharyya, D.; Tao, W.; Berman, A.; Broxmeyer, H. E.; Keenan, R. J.; Glick, B. S. A noncytotoxic DsRed variant for whole-cell labeling. *Nat. Methods* **2008**, *5*, 955-957.
- (252) Whitaker, M. Genetically encoded probes for measurement of intracellular calcium. *Methods Cell Biol.* **2010**, *99*, 153-182.
- (253) Campanella, J.; Bitincka, L.; Smalley, J. MatGAT: An application that generates similarity/identity matrices using protein or DNA sequences. *BMC Bioinformatics* **2003**, *4*, 29.

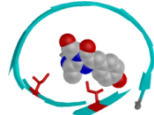
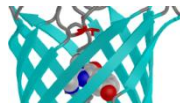
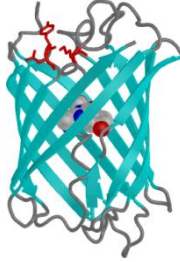
Appendices

Appendix A: Supplementary data for Chapter 2

Table A1 Definitions of variants described in Chapter 2.

Protein variant	Substitutions and modifications relative to parent
mTFP1 ^a	= cFP484 ^b + delete all residues before 6a and add MVSKGEE, H42N, L44I, S62T, N63T, Q66A, L72F, A80P, D81N, R123H, F124L, D125K, M127E, L141T, K142G, E144D, P145A, I149R, V158K, I161V, S162K, S164K, Y173H, C175V, S179T, K182R, V186A, L213V, N216S, Y221N, L223T, L224D, delete all residues after 224 and add GMDELYK.
mClavGR1 ^c	= mTFP1 + G6dS, K17R, E34S, Y38F, D39E, N42Q, N45D, S57A, A66H, A71V, N81D, T96S, K102G, V105C, K106I, V107A, K108T, S109N, S112T, E115K, Y120N, E121K, L124F, D144E, R149K, H163M, E168K, H173Y, V175C, K178R, I180T, R182K, Q184a insertion, N203S, T212K, V213L, S216H, R220H, N221S, S222G, T223L, D224P
mClavGR1.1	= mClavGR1 + T6bl, S34R, K36R, K74R, M113I, L166Q, Y173H
dClavGR1.6	= mClavGR1 + T6bl, T41I, D77E, F99Y, K115E, E127V, K139R, A183V
mClavGR1.8	= mClavGR1 + T6bl, T41I, D77E, F99Y, K115E, K139R, A183V
mClavGR2	= mClavGR1 + T6bl, T41I, D77E, F99Y, K115E, E127T, K139R, A183V
^a Previously described [92]. ^b cFP484 is the wild-type protein from <i>Clavularia</i> (Genbank accession AAF03374) . ^c The key monomerizing mutations R123H, D125K, and M127E (of the A-B interface) and S162K and S164K (of the AC interface) of mTFP1 have been retained in mClavGR1.	

Table A2 Positions targeted for partial or complete randomization by mutagenesis. Figures were generated from PDB ID 2HQK [92] using MacPyMol (DeLano Scientific) [253]. Residues being substituted are indicated in red.

Template	Residue(s)	Substitution(s)	Codon(s)	Result	Notes
mClavGR1	I44	All 20 amino acids	NNK	no improved variants	 Close proximity to chromophore. Substitutions at residues 44 and 199 previously shown to benefit KikGR [170].
	I199	Leu, Ile, Met, Val, His, Gln, Asn, Lys, Asp, Glu	VWK		
mClavGR1	T73	All 20 amino acids	NNK	no improved variants	 Previously reported to influence pK_a of related FPs [93].
dClavGR1.6	F38	All 20 amino acids	NNK	no improved variants	 A cluster of residues located near the 'top' of the central helix. Substitution of residue 74 previously shown to benefit KikGR [170].
	E39	Asp, Glu, Gly, His, Lys, Asn, Gln, Arg, Ser	VRN		
	K74	Asp, Glu, Gly, His, Lys, Asn, Gln, Arg, Ser	VRN		




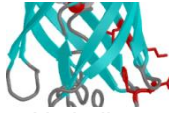
mClavGR1.6	E127V	Glu, Thr	GAA (Glu) ACB (Thr)	identified E127T	 Disrupted residual dimerization tendency
dClavGR1.6	K162	All 20 amino acids	NNK	no improved variants	 Adjacent to M163, which is in close proximity to the chromophore. K162A similar with K162K.
dClavGR1.6	M163	His or all 20 amino acids	CAY or NNK	no improved variants	 Interacts with phenolate of chromophore. M163K brighter in green state but did not photoconvert to the red state.
mClavGR2	F124	Phe, Ile, Leu, Val	NTT	no improved variants	 K125N similar to K125K
	K125	His, Lys, Gln, Asn	AAM, MAG		
	N128	Pro, Arg, Ser, Thr	MSC		
	F129	Phe, Ile, Leu, Val	NTT		

Table A3 Amino acid similarities (lower triangle) and identities (upper triangle) as calculated using MatGat [253] using the complete protein sequences provided in Figure 2.2.

Protein	mTFP1	mClavGR1	mClavGR2	mEos2	Dendra2	Kaede	mKikGR
mTFP1	--	81.9	79.3	62.4	66.9	59.1	59.3
mClavGR1	90.7	--	96.6	77.6	70.2	75.1	71.5
mClavGR2	89	98.3	--	76.4	69.3	73.8	71.1
mEos2	76.3	84.4	84.4	--	70.9	83.2	71.7
Dendra2	78	81	80.6	84.1	--	67.7	64.4
Kaede	76.3	84	83.5	91.6	82.7	--	66.1
mKikGR	77.5	84.8	84.4	83.6	80.2	80.6	--

Table A4 Amino acid similarities (lower triangle) and identities (upper triangle) as calculated using MatGat [253] using only the sequence from position 6c to 224 in Figure 2.2.

Protein	mTFP1	mClavGR1	mClavGR2	mEos2	Dendra2	Kaede	mKikGR
mTFP1	--	80.5	78.3	66.5	71.8	62.4	63.2
mClavGR1	90	--	96.8	82.8	74.2	79.6	76.3
mClavGR2	88.7	98.6	--	81.4	73.3	78.3	75.9
mEos2	80.5	89.6	89.6	--	71	84.2	74.1
Dendra2	83.2	85.1	84.6	83.7	--	67.9	65.9
Kaede	79.6	88.7	88.2	92.3	82.4	--	68.3
mKikGR	80.7	88.8	88.3	86.1	82.1	83.4	--

Table A5 Amino acid differences between green-to-red proteins, considering only those residues with side chains directed towards the interior of the β -barrel. Amino acids that differ from mClavGR2 are underlined.

Position	mClavGR2	mEos2	Dendra2	Kaede	mKikGR
14	I	I	<u>V</u>	I	I
16	L	L	<u>V</u>	L	L
44	I	<u>M</u>	<u>A</u>	<u>M</u>	<u>V</u>
52	A	<u>G</u>	A	A	<u>G</u>
57	A	A	<u>S</u>	A	A
58	Y	<u>F</u>	Y	Y	<u>F</u>
65	F	F	<u>V</u>	F	F
73	T	<u>A</u>	T	<u>A</u>	<u>V</u>
91	Y	Y	Y	<u>F</u>	Y
97	M	<u>L</u>	M	<u>L</u>	M
99	Y	<u>F</u>	<u>F</u>	<u>F</u>	Y
107	A	A	<u>I</u>	A	A
109	N	N	<u>S</u>	N	N
120	N	N	<u>Q</u>	N	N
122	I	<u>V</u>	<u>V</u>	<u>V</u>	I
150	M	M	<u>L</u>	M	M
161	V	<u>I</u>	<u>I</u>	<u>I</u>	V
183	V	<u>A</u>	<u>A</u>	<u>S</u>	<u>A</u>
199	I	I	I	I	<u>M</u>
213	L	L	L	L	<u>A</u>
# differences from mClavGR2		9	13	8	7

Table A6 Sum of the number of differences between green-to-red proteins, considering only those residues with side chains directed towards the interior of the β -barrel.

	mEos2	Dendra2	Kaede	mKikGR
mClavGR2	9	13	8	7
mEos2	--	13	4	8
Dendra2	--	--	13	17
Kaede	--	--	--	12

Appendix B: Supplementary data for Chapter 3

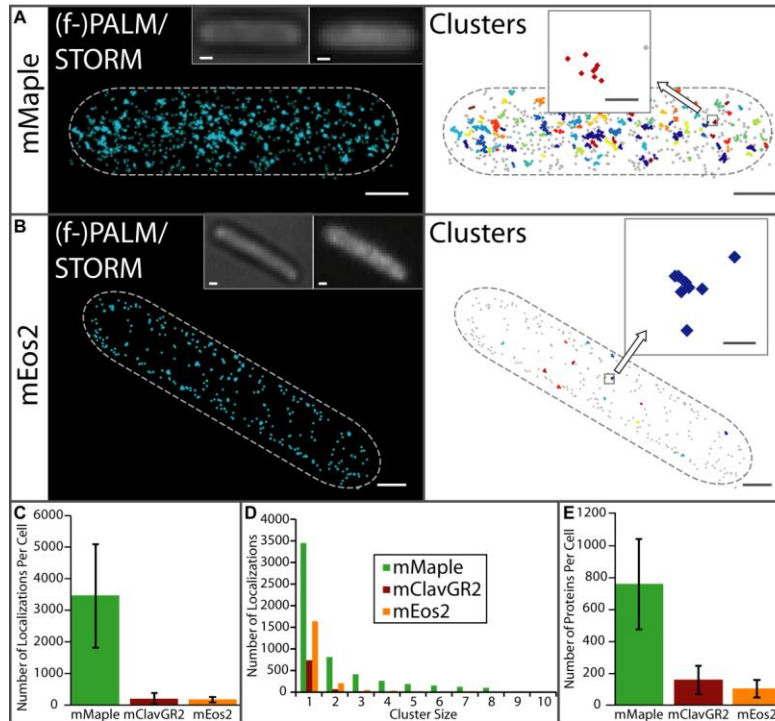


Figure B1 (f-)PALM/STORM characterization of the number of observed localizations and proteins per cell. (A-B) (f-)PALM/STORM images of fixed *E. coli* expressing cytoplasmic (A) mMApple ($N = 1696$ localizations), or (B) mEos2 ($N = 472$ localizations). Localizations are represented as normalized 2D Gaussian peaks with widths given by their theoretical localization precisions (left panels) and plotted as small markers grouped into clusters with adjacent spacing of 30 nm or less (right panels). Individual protein localizations are shown in grey whereas closely spaced localizations (< 30 nm) are grouped into clusters of the same color (right panels). The bright field and conventional fluorescence images are shown for comparison (left panels, left and right inset respectively). Scale bars are 500 nm. (C) Average number of localizations per cell for each cytoplasmically expressed pcFP. (D) The distribution of cluster sizes (< 30 nm interlocalization spacing) for cytoplasmically expressed pcFPs. (E) Average number of cytoplasmically expressed proteins per cell. Rather than counting each localization as a single molecule, we count each cluster of localizations (localizations spaced < 30 nm) as a single protein. The dotted lines in (A-B) denote the *E. coli* cell boundary. Scale bars are 500 nm and 50 nm (zooms). Zooms in (A-B) show possible reversible photoswitching events of single proteins. Error is the standard deviation ($N = 20$ cells (mMaple), $N = 17$ cells (mClavGR2), $N = 16$ cells (mEos2)). The large error bars are primarily due to variation in protein expression

between cells. Experiment and analysis described in this figure is performed by Ann McEvoy.

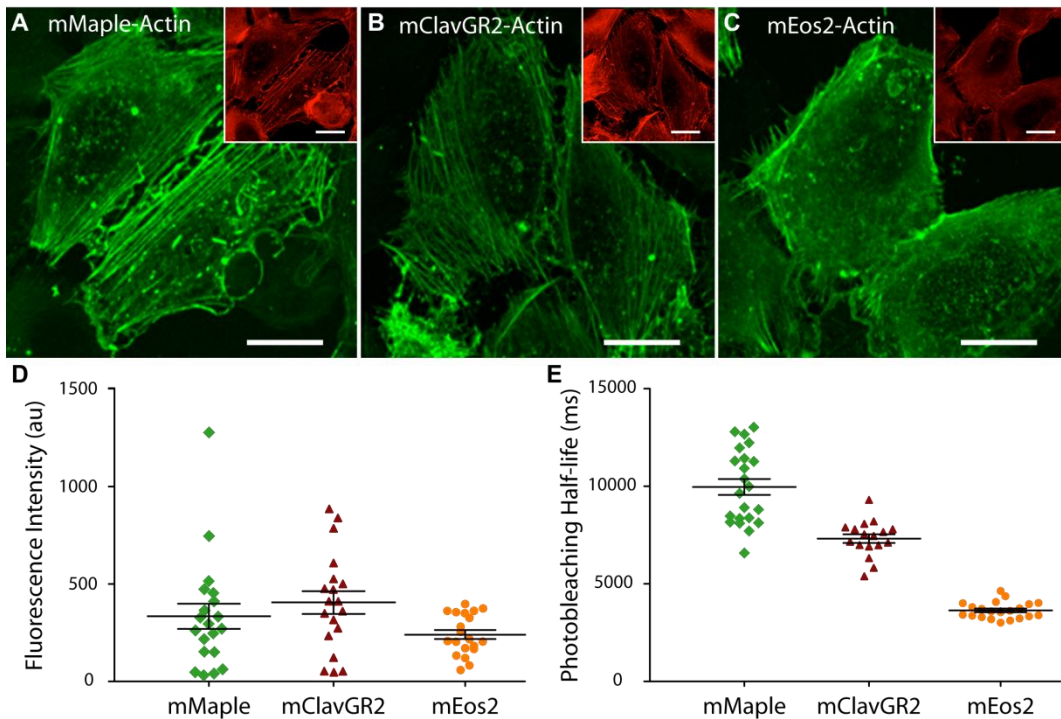


Figure B2 Photostability and green-state brightness characterization of pcFP-actin fusions in mammalian cells. Representative widefield fluorescence images of U2OS cells transfected with plasmids encoding either mMaple-actin (A), mClavGR2-actin (B), or mEos2-actin (C). The cells were stained with phalloidin as a comparison (insets in A-C). (D) All pcFP-actin fusions have similar cellular intensities. However, mMaple is more photostable than the other two pcFP fusions (E). The data in (E) were found to be statistically significant using the Mann-Whitney t-test ($p \leq 0.0001$). Experiment and analysis described in this figure is performed by Evgenia Platonova from the Helge Ewers group at ETH Zurich.

Appendix C: Supplementary data for Chapter 4

Table C1 Genealogy and mutations in GR-GECOs

Protein	Mutations	Notes
GR-GECO0.1	cpmMaple (145Y-143W) + M68I/ N141H/ G145S/ K284M	Replace the cpGFP module with cpmMaple in G-GECO1.1. The result is a 3-part chimera composed of M13, cpmMaple and calmodulin.
GR-GECO0.2	GR-GECO0.1 + 62G/ 63Y/ V73aE/ D269N/ 302E/ 302aP	Best variant found in the termini libraries
GR-GECO0.3	GR-GECO0.2 + E61P/ N380W/ D381G	Best variant found in and semi- saturation libraries at position E61, R377, N380 and D381
GR-GECO0.5	GR-GECO0.3 + L60P/ K126E/ M146V/ T165A/ E197K/ Q244R/ R304H/ V430E/ N440S	Intermediate variant found during error-prone libraries and StEP libraries
GR-GECO1.1	GR-GECO0.5 + S88G/ T143A/ M158I/ A165T/ S276T/ T303V/ H304T/ K333E/ T413A/ T420A/ M427I	Variant displays good Ca ²⁺ response at both green and red states in saturation libraries at position T303/R304
GR-GECO1.2	GR-GECO0.5 + S88G/ A165T/ I202V/ G262S/ T413A/ K418R/ E430V/ D434N	Variant displays best Ca ²⁺ response at the red states in saturation libraries at position T303/R304

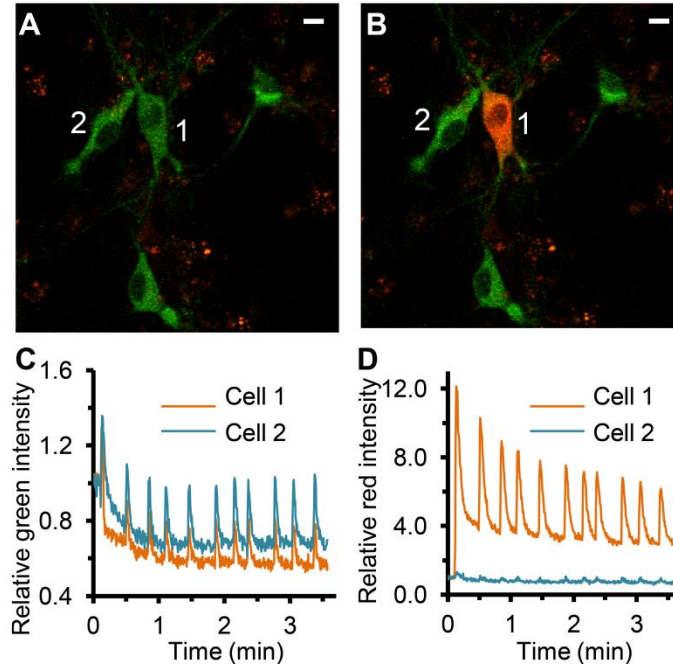


Figure C1 GR-GECO1.2 expressed in dissociated hippocampal neurons. (A, B) Merged green and red fluorescence images of cells before (A) and after (B) photoactivation of cell 1. Scale bar 10 μm. (C, D) Imaging of spontaneous Ca^{2+} oscillations in the green (C) and red (D) emission channels. Only the highlighted cell 1 was observed in the red channel. Experiment and data analysis are credited to Tomoki Matsuda.

Appendix D: Supplementary data for Chapter 5

Table D1 Positions targeted for mutagenesis

Template	Target(s)	Result	Notes
tPapaya0.01	I106R/K	I106R	A-B interface broken, dim.
dPapaya0.1	D164K/R R178D/N/K/H/Q/E	D164K, R178H	A-C interface disrupted, extremely dim.
mPapaya0.27	Delete F226-A231 or replaced with "RNSTD"	Δ F226-A231	Dimeric tendency in mPapaya0.27 significantly reduced.
mPapaya0.3	C149S/T	C149T	Dimeric tendency reduced.
mPapaya0.4	Y127R	Y127R	Dimeric tendency reduced.
mPapaya0.5	S166K/R	S166K	Dimeric tendency abolished.